50X1-HUM

Next 2 Page(s) In Document Denied

BOOK

Terrestrical Magnetism

(UNCL) ~~MAGNETIC VARIATIONS AND POLAR AURORA (MAGNITNYYE VARIATSII I POLYARNYYE SIYANIYA)~~ chapt. VI - VII

AUTHOR: B.M. YANOVSKII

SOURCE: ~~TERRESTRIAL MAGNETISM~~ Zemnoy magnetizm

~~1953~~ pp. ~~181-231~~

1953

pp. 181 - 273

STAT

CHAPTER VI

MAGNETIC VARIATIONS AND THE AURORA BOREALIS.

Section 1. The Solar-Diurnal Variations.

The solar-diurnal variations, which we shall denote by the letter S, consist of periodic variations of the elements of terrestrial magnetism with a period equal to the length of the solar day.

A characteristic feature of these variations is their occurrence according to local time. For this reason, at two different longitudes, the phases of the fluctuations of one element or another will differ by the difference in longitudes between the two points. Thus, if we represent the deviation from its mean value, i.e. the variation, of any element at a given point of the earth's surface in the form of a simple harmonic oscillation:

$$S = S_0 \sin \frac{2\pi t}{T},$$

where S_0 is the amplitude, T are the solar days and t the local time, then the variation at another point whose longitude differs by λ, is represented by the equation:

$$S' = S_0 \sin \frac{2\pi}{T}(t + \lambda).$$

Fig. 68 represents the mean annual diurnal march of the declination, the horizontal and vertical components, i.e. the relation of the variations of these elements and the local time during a 24 hour period, according to the observations of

the Pavlovsk magnetic observatory.

The mean annual values of the variations of the various elements respectively are plotted along the axis of ordinates and the local time along the axis of abscissas.

On considering the curve for the variations of the declination, it will be seen that the magnetic needle which remains quiet at night (curve δ D), is deflected in the morning towards the east, and by 0800 hours reaches the maximum deflection, after which it begins to move in the opposite sense, and by 1400 hours it reaches its maximum deviation towards the west. The remaining curves show that the horizontal component (the curve δ H) has a minimum at about 1100 hours and a maximum about 2000 hours, while the vertical component, which remains almost unchanged during the night, begins in the morning to increase, and after noon reaches its maximum value. As shown by observations, the diurnal march of the elements of terrestrial magnetism does not remain constant but varies irregularly from one day to another; in this case the amplitude of fluctuations are mainly subject to change, while the phases themselves remain almost unchanged.

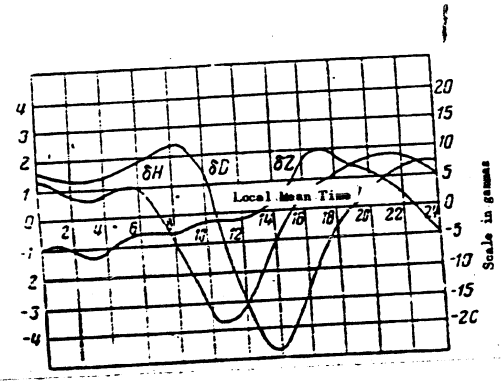

Fig.68. Diurnal March of the Elements of Terrestrial Magnetism at the Pavlovsk Magnetic Observatory.

Table 18 gives the values of the differences between the maximum and minimum values of each element at different seasons at Pavlovsk.

Table 18

Season	Difference between the maximum and minimum		
	D	H	Z
Winter	4.1′	09γ	7γ
Spring	7.9	27	12
Summer	12.0	44	20
Autumn	8.3	36	12

The Table shows that the variations in the diurnal march increase from the winter months, when the declination of the Sun is smallest (to the summer months, when the declination of the Sun is greatest (+ 23,5°).

The next feature of the solar-diurnal variations is their dependence on the value of the magnetic activity on one day or another. For this reason two forms of solar diurnal variations are distinguished: The variations on quiet days when which are obtained by working up the observations only for quiet days only, and variations in stormy days, which are called disturbed variations and are obtained by working up the observations on stormy days. The former are denoted by S_d, the latter by S_q.

Variations on stormy days differ markedly from the variations on quiet days. This difference particularly effects the march of the variations of the vertical component, where not only the amplitude but the whole character of the curve changes. In addition, the amplitudes of the quiet diurnal variations S_q vary during the course of the year, taking their maximum value during the summer solstice and their minimum values during the winter solstice. During the epoch of the equinoxes, the amplitude is the mean between the winter and summer and is the same in both hemispheres.

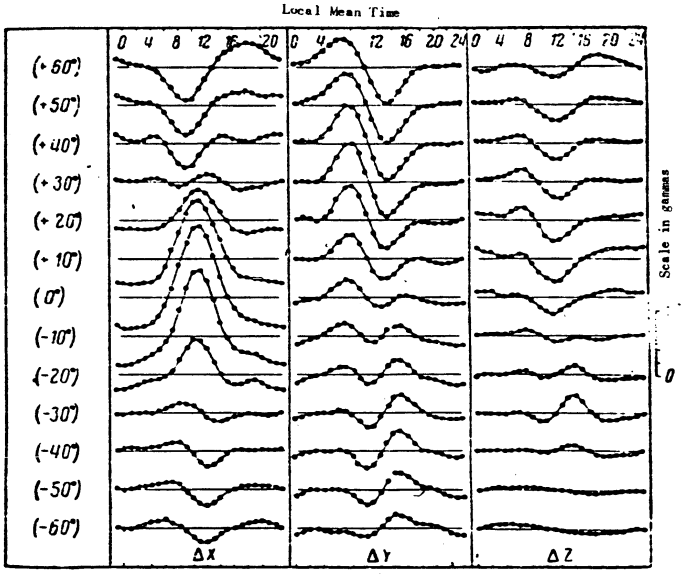

Fig.69,a. Diurnal March of the Elements X, Y and Z on Quiet Days During the Period of the Summer Solstice at Various Latitudes.

Moreover, observations show that the solar diurnal variation at various points of the earth's surface is of different character. For points located on one and the same parallel, however, the diurnal march is almost the same, but for points located along a meridian, it varies according to a certain definite law. Fig.69 represents the curves of the diurnal march on quiet days at different latitudes for the three elements X, Y and Z during the period of the summer solstice in the northern hemisphere (Fig.69,a), for the winter solstice (Fig.69,b), and for the epoch of the equi-

nox (Fig.69,c), both spring and autumn. These curves show that the variations of the northern component δX have approximately the same character in northern and southern latitudes, since for the individual elements in the southern latitudes, they are inverted on passage across the magnetic equator, i.e. they are mirror pictures of the variations in the northern latitudes.

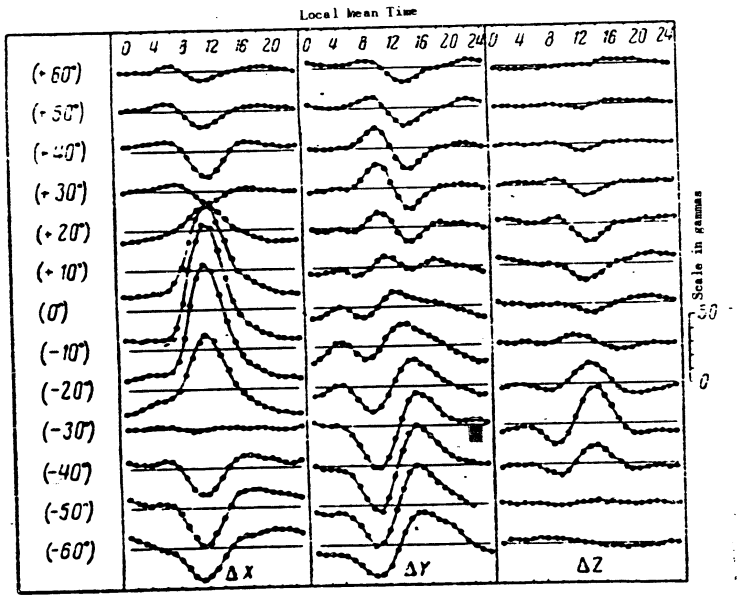

Fig.69,b. Diurnal March of the Elements X, Y and Z on Quiet Days During the Periods of the Winter Solstice at Various Latitudes.

On the equator itself, the variations on the eastern and vertical components are close to zero. For the northern component, such a reversibility of the curves takes place at magnetic latitude of about 30° in both southern and northern hemispheres.

The variation of the amplitudes during the course of the year will be clearly seen by comparing Figures 69,a, 69,b, and 69,c.

Fig.70 gives analogous curves of the diurnal march of the variations on stormy days S_d for various latitudes. The ordinates of these curves represent the difference between the variations on stormy days and variations on quiet days ($S_d - S_q$). These curves show that in the low latitudes the differences $S_d - S_q$ are small and, consequently the S_q variations predominate in them, while in the high latitudes, on

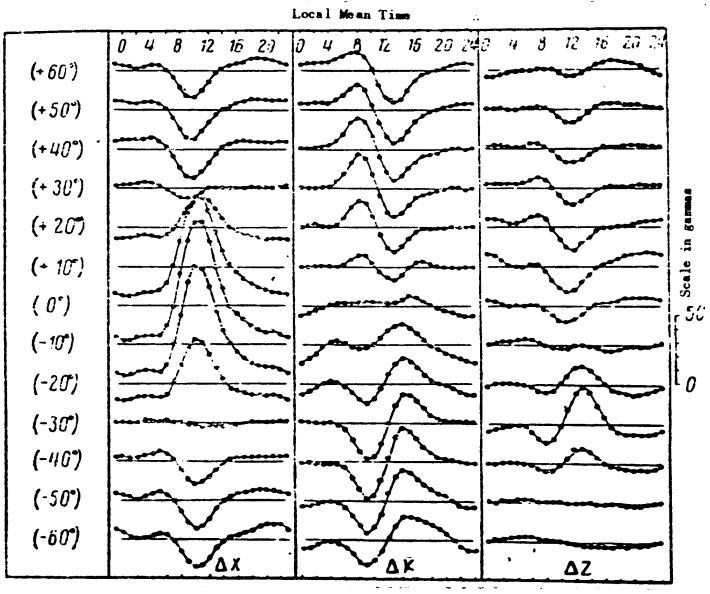

Fig.69,c. Diurnal March of the Elements X, Y and Z on Quiet Days in the Periods of the Equinoxes at Various Latitudes.

the contrary, the variations $S_D = S_d - S_q$ play a predominant part.

The most striking idea of the march of the diurnal variations given by the construction of the so-called vector diagrams which represent the projection of the vector variations S_q on the horizontal and vertical planes.

Such a diagram of the projections of S_q of the horizontal plane is shown in Fig.71 in the period of the equinox for a portion of the earth's surface turned

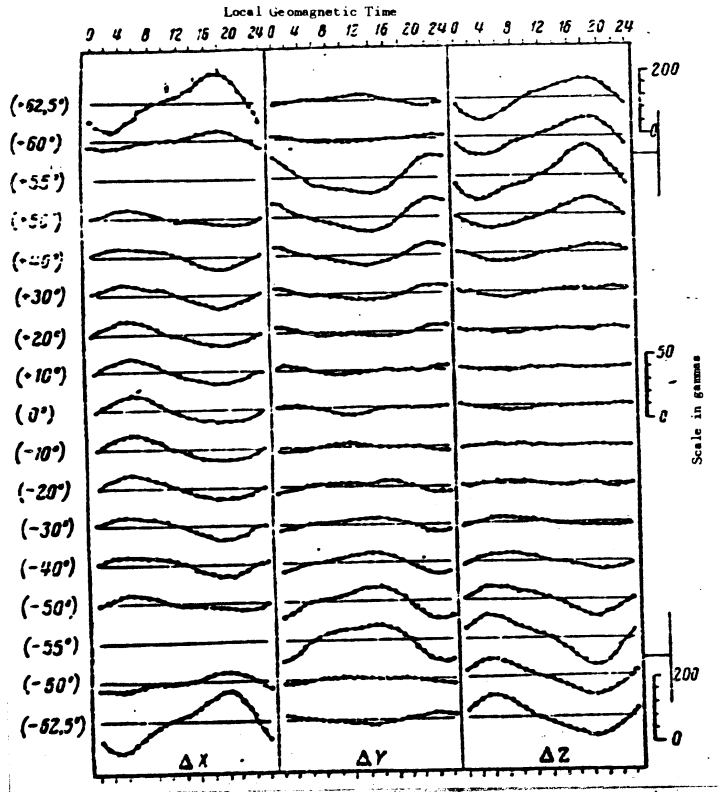

Fig.70. Diurnal March of Variations on Stormy Days at Various Latitudes.

towards the sun and bounded by the geographical latitudes from + 60 to - 60° and the longitudes from 6 to 18 hours. In this case the longitudes coincide with the local time from 6 to 18 hours.

The following may be noted from this diagram: the vector S_q in the northern hemisphere during the daylight hours is always directed towards a certain center located on the forenoon meridian at the parallel + 30°, while in the southern hemisphere it is directed from a center located on the same meridian and the parallel -30°.

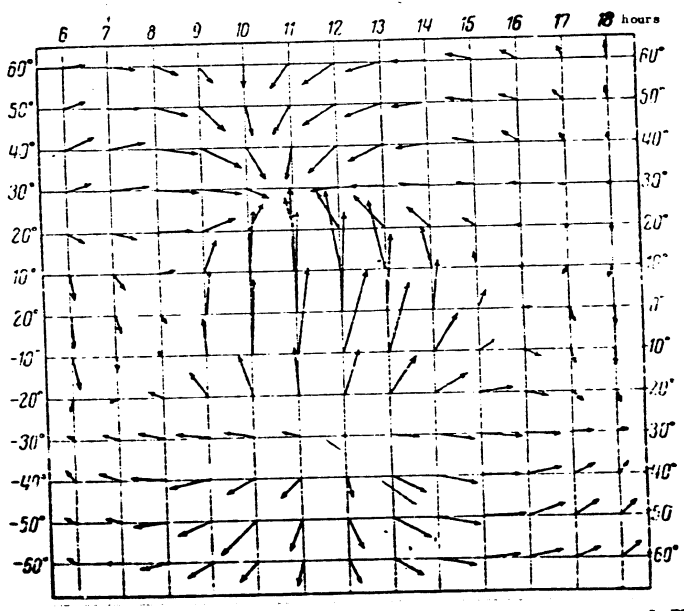

Fig.71. Projection of the Vector of Variations S_q on the Horizontal Plane.

Fig.72,a and 72,b show the diagrams of projections of the vector S_q on the vertical planes: one of them on the plane of the principal meridian, and the other on the plane of a great circle making contact with the 30th parallel at the meridian point. Both diagrams show that the centers towards which the S_q vectors are directed lie above the earth's surface roughly above the parallels + 30 and - 30 near the principal meridian.

Formally, these centers may be identified with the axis of the eddy current whose sense is counterclockwise, viewed from above, in the northern hemisphere, and clockwise in the southern hemisphere. Thus the diurnal variations may be explained by the existence in the atmosphere of a system of closed eddy currents which remain fixed in space, and within which the earth rotates. Since the maximum value of the vector of the variations come during the daylight hours, the maximum current strength must be in the space between the sun and the earth.

The system of electrical currents corresponding to the fields of diurnal variations. The magnetic field of the solar diurnal variations, i.e. the field corresponding to the distribution of the vector of variations must have its sources which most probably may consist of a certain system of electrical currents which most probably may be represented in the form of a certain system of electrical currents. The general distribution of the vector of variations on the earth's surface for a given moment of time, represented in Fig.72,a and 72,b, indicates that the system of currents with their center at latitude 30° and on the principal meridian, remains fixed in space between the sun and the earth, and an observer on the earth, in rotating with respect to this system, passes during the course of a 24 hour day through all values of the vector of field strength of these currents, distributed along the parallel of the observer. Since the electric current and its magnetic field are connected by the Biot-Sawra law, then, if we know the field, the current may also be determined, provided that the distances between the current and those points at which the field is known are also known. For this reason to find the sys-

tem of currents according to an assigned distribution of a magnetic field, we must start out from the law in question or from its consequences.

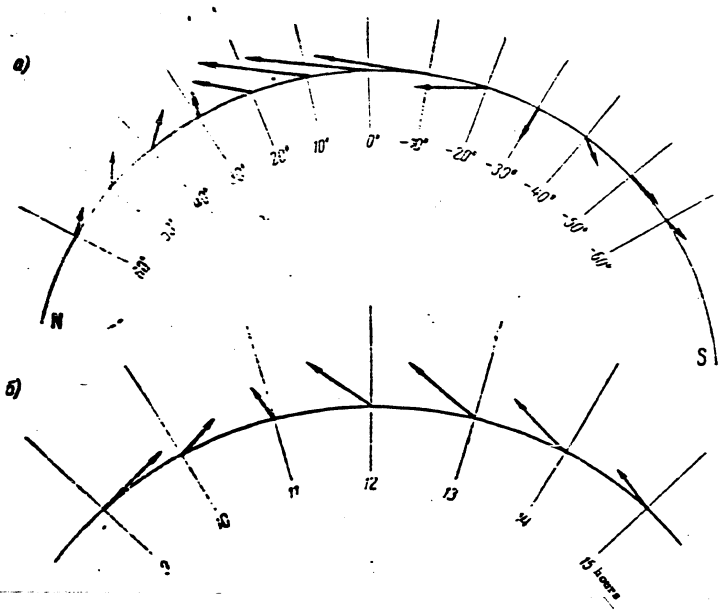

Fig.72. Projection of the Vector of Variation S_q on to the Vertical Planes; a. Onto the Plane of the Principal Meridian; b. Onto the Plane of the Prime Vertical.

Such a consequence, which is the most convenient to calculate, is the equivalence of a closed current with a dual magnetic layer. For this reason, by replacing the system of currents by a dual magnetic layer of variable density, disposed on a sphere concentric with the earth's surface and having a radius R greater than the earth's R radius, we find the magnetic potential U of such a layer at the earth's surface. We know from the theory of potential that

$$U = \int\int \mu \frac{\partial}{\partial n}\left(\frac{1}{\rho}\right) dS,$$

where μ is the surface density of the magnetic moment of a double layer corresponding to the element of surface $d\sigma$; ρ is the distance between the point of the earth's surface where the potential U is sought; n is the direction of the normal to the surface of the sphere coinciding with the redius R, and the integration is taken over the entire surface of the sphere.

Since $\rho^2 = R^2 + r^2 - 2Rr\cos\gamma$, where γ is the angle between R and r, and $r > R$, then

$$\frac{1}{\rho} = \sum_{n=0}^{\infty} \frac{R^n}{r^{n+1}} P_n(\cos\gamma)$$

and, consequently,

$$U = \int\int \mu \frac{\partial}{\partial n} \sum_{n=1}^{\infty} \frac{R^n}{r^{n+1}} P_n(\cos\gamma)\, dS,$$

or

$$U = -\sum_{n=1}^{\infty} (n+1) \frac{R^n}{r^{n+2}} \int\int \mu P_n(\cos\gamma)\, dS.$$

It is proved in potential theory that every function of two variables (the latitude θ and longitude λ), assigned for the points of a sphere, may be uniquely expanded into a series in Laplace spherical conditions; and for this reason μ may be represented in the form of the sum:

$$\mu = \sum_{k=1}^{\infty} \mu_k, \quad \text{где} \quad \mu_k = \sum_{m=0}^{k} (a_n^m \cos m\lambda + b_n^m \sin m\lambda) P_n^m(\cos\theta). \quad (6.1)$$

Consequently,

$$U = -\sum_{n=1}^{\infty}(n+1)\frac{R^n}{r^{n+2}}\sum_{k=1}^{\infty}\int\int \mu_k P_n(\cos\gamma)\,dS.$$

It is known from the theory of spherical functions that

$$\int\int \mu_k P_n(\cos\gamma)\,dS = \frac{4\pi r^2}{2n+1}\mu_n \quad \text{если} \quad k = n,$$
$$\int\int \mu_k P_n(\cos\gamma)\,dS = 0, \quad \text{если} \quad k \neq n.$$

For this reason

$$U = -\sum_{n=1}^{\infty} 4\pi\frac{n+1}{2n+1}\left(\frac{R}{r}\right)^n \mu_n. \qquad (6.2)$$

On the other hand, the magnetic potential of the field of variations, as we know, is expressed in the form of a series:

$$U = R\sum U_n, \quad \text{где} \quad U_n = \sum (p_n^m \cos m\lambda + q_n^m \sin m\lambda) P_n^m(\cos\vartheta), \qquad (6.3)$$

where the coefficients p_n^m and q_n^m are known from observations and represents the external part of the potential of variations.

On comparing the expressions 6.2 and 6.3, we find that

$$RU_n = -4\pi\frac{n+1}{2n+1}\left(\frac{R}{r}\right)^n \mu_n,$$

whence

$$\mu_n = -\frac{R}{4\pi}\cdot\frac{2n+1}{n+1}\left(\frac{r}{R}\right)^n U_n,$$

and, by substituting μ_n in equation 6.1 and replacing U_n by its value, we have

$$\mu = -\frac{R}{4\pi}\sum_{n=1}^{\infty}\frac{2n+1}{n+1}\left(\frac{r}{R}\right)^n \sum_{m=0}^{n}(p_n^m \cos m\lambda + q_n^m \sin m\lambda) P_n^m(\cos\vartheta). \qquad (6.4)$$

The density of the magnetic moment μ of the double layer is equivalent to the current strength of a closed circuit, and for this reason the expression for the strength of the current at any point of the surface of the sphere with coordinates θ and λ, will have the same form as equation 6.4.

This equation shows that for the calculation of the current strength at any point of a sphere it is sufficient if we merely multiply each term of the expansion of the potential of order n, by the term containing one unknown quantity, the radius of the external sphere through which the current flows. This radius must be determined from any other considerations. In the first paper devoted to the calculations of the currents in the atmosphere, this radius was fortuitously taken at 100 km more than the earth's radius, which corresponds very closely to present-day data on the height of the conducting layer in which the existence of such currents is possible.

At the present time there is ground for holding that the system of currents causing the solar diurnal variations of the magnetic field is located in the E layer of the ionosphere, the height of which, according to observations on deflection of radio waves, varies between 100 and 120 km. For this reason, if we assume the height of the E layer as equal to 109 km, as has been done in Ben'kova's paper and t is expressed in thousand amperes, while p_n^m and q_n^m are expressed in gammas, then equation 6.4 takes the form:

$$I = -5{,}0707 \sum_{n=1}^{\infty} \frac{2n+1}{n+1} (1{,}02)^n \sum_{m=0}^{n} (p_n^m \cos m\lambda + q_n^m \sin m\lambda) P_n^m (\cos\theta). \quad (6.5)$$

It is this formula that will serve for the calculation of the system of current. On determining from it the value of I for various values of λ and θ, and plotting these values on a map, we may draw a series of isolines (lines of equal current strength), which will represent current lines, i.e. the lines along which the current

flows. The difference between the values of the current strength on two adjacent isolines gives us the value of the current flowing between these lines.

Such a system of currents, correspondent to the solar diurnal variations in the equinoctial period and for heights of 100 km, is shown in Fig.73, while Fig.74 shows the system of currents corresponding to the summer solstice.

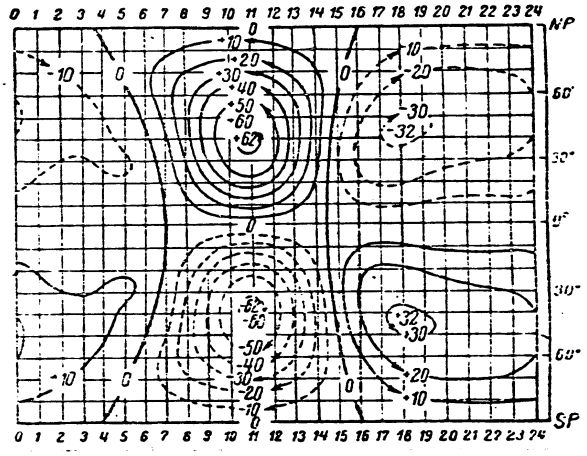

Fig.73. System of Electric Currents Corresponding to the Solar Diurnal Variations in the Equinoctial Periods.

The closed curves showing the direction of the currents are drawn in such a way that a current of 1000 amp. flows between two adjacent curves. The currents flow in four main systems of circuits, two northern and two southern. In this case, two systems of contours are located on the lighted hemisphere and the two others on the night hemisphere, the former being more intense.

The total current in the daylight circuit is equal to 62,000 amp. during the time of the equinox and 89,000 amp. during the time of the solstices.

As stated in Section 3, about a third of the field of variations is due to internal causes, i.e. to currents flowing within the earth. It may be assumed that these currents are caused by induction of the magnetic field of the external currents in some conducting layer of the earth or perhaps in the entire earth. If the conductivity of the earth were known, then the determination of the system of currents would reduce down to the above mentioned operation, analogous to the operation of calculating the external currents. But we know nothing at all of the conductivity of the interior parts of the earth. We know only the conductivity of the upper stone envelope, which is of the order of 10^{-6} ohm^{-1} cm^{-1} and the conductivity of the oceans, of the order of 4×10^{-2} ohm^{-1} cm^{-1}.

For this reason there have been attempts to calculate the induction current and the magnetic field caused by them under the assumption that the conductivity of the earth is everywhere the same.

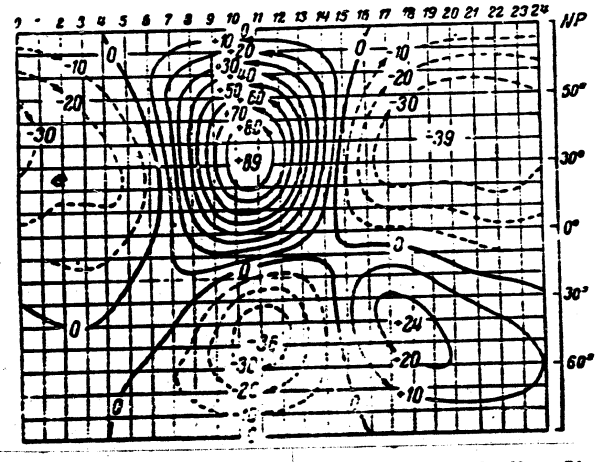

Fig.74. System of Electric Currents Corresponding to the Mean Diurnal Variations During the Period of the Summer Solstice.

It was found that to make the calculated values of the variation agree with the observed values, it would be necessary to assume a conductivity of $\gamma = 3.6 \times 10^{-4}$ ohm^{-1} cm^{-1} for the earth, that is, somewhat less than sea water and more than the conductivity of the upper layers. In addition, it must also be recognized that the upper layer is probably nonconducting down to a depth of 300 km.

The system of currents shown in Fig.73 and 74 has been calculated by equation 6.5 under the assumption that the solar diurnal variations are functions of the geographical latitude at a local time. But a comparison of the observed curves of the diurnal march of the S_q variations with the curves calculated by equation 6.5 in which the arguments are the geographical latitude and the local time, does not yield good agreement. For this reason the S_q variations were expanded into spherical harmonics, taking the geomagnetic latitude and the geomagnetic time as the independent variables (Bibl.16). The agreement between and observed values was now considerably better. Fig.75 shows diurnal march of the northern component at the Huancayo Observatory (South America, $\varphi = -12.0^\circ$, $\lambda = 284.7^\circ$), observed and calculated by equation 6.5 and by formulas in which φ and λ are geomagnetic. As will be seen the latter curve is in considerably better agreement with the observed curve than the former one.

Fig.76 shows a system of currents corresponding to the expansion in geomagnetic coordinates. Its principal difference from system in Fig.73 is the asymmetry of the currents of the northern and southern hemisphere; the current in the southern hemisphere is two and a half times as great as that in the northern hemisphere. In addition, their centers are also asymmetric with respect to the equator. It is possible to explain this by the fact that the magnetic field is symmetric with respect to the magnetic axis not the geographic axis, which should lead, according to the dynamo theory, to a reduction of current in the northern hemisphere and its weakening in the southern hemisphere (sic). But there are no quantitative calculations.

For a greater approximation of the observed diurnal march to the theoretical,

N.I. Ben'kova (Bibl.47) made a spherical analysis of the diurnal variations of S_q at

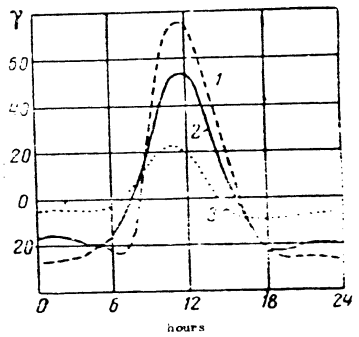

Fig.75. Diurnal March of Northern Component at Huancayo Observatory.
1. observed curve; 2. calculated in geomagnetic coordinates; 3. calculated in geographic coordinates
abscissa: hours.

47 observatories, including some beyond the Arctic Circle, allowing for the fact that the variation depends not only on the latitude but also on the longitude of the place. In this case the longitude, latitude and time taken were all geomagnetic. No one before Ben'kova has ever performed the expansion with such a formulation of the question, since all of the investigators had considered the diurnal variations to be independent of the longitude.

Since the diurnal analysis assumes the expansion of a function depending only on two coordinates, Ben'kova assumed, in order to calculate the third coordinate, longitude, that the S_q variations are the sum of two functions, one of which, S_{q1},

depends on the latitude Φ and the local time t, while the secone one, S_{q2}, depends on the longitude Λ and the latitude, i.e.

$$S_q = S_{q1}(\Phi, t) + S_{q2}(\Phi, \Lambda), \qquad (6.6)$$

and she performed the expansion separately for the functions S_{q1} and S_{q2}.

The determination of the coefficients of the expansion was performed by the method set forth in Section 3, by equation 5.4 and 5.8. To eliminate the influence

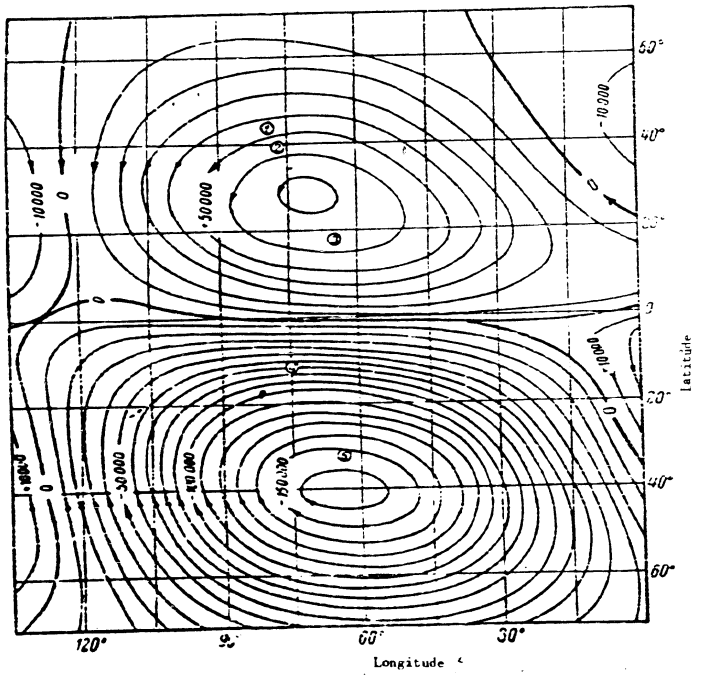

Fig.76. System of Electrical Currents According to Calculations of Mac Nish.

of the longitude in the expressions for S_{q1}, the means for the given latitude were taken instead of the coefficients calculated for the separate observatory as the initial coefficients a_x^m etc. in equation 5.8.

The original material for the functions S_{q2} was the differences between the mean values of the coefficients a_x^m and those calculated by equation 5.4. As a result of this analysis, performed for a large number of stations, it was found that the main part of the field of variations is represented by the function S_{q1}, i.e. it does not depend on the longitude.

The longitude function exists, but its influence is shown only at low and middle latitudes, so that in the polar regions the introduction of longitude terms does not improve the agreement between the calculated and observed values of the variation.

In accordance with the result of this analysis, Ben'kova constructed a system of points for the summer months (May-August) both for the functions S_{q1}, and for the function S_{q2}. They are shown in Fig.77 and Fig.78. A comparison with the current maps on Fig.76 shows that they are in better agreement with the maps of Fig.73 than with the maps of Fig.76. The eddy of current located on the daylight side of the northern hemisphere has about the same form as in Fig.73, and is situated at the

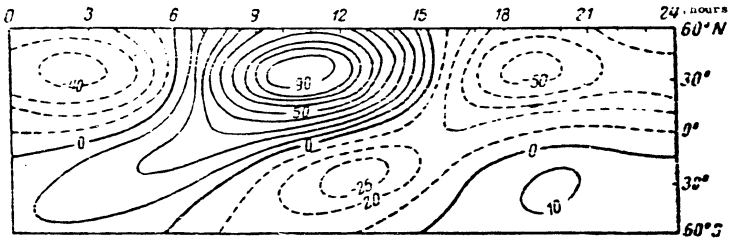

Fig.77. System of Currents for the Summer Months, from Calculations of N. P. Ben'kova (for functions S_{q1}).

same latitude, but Ben'kova found the intensity of this eddy to be greater than that of the eddy in Fig.73. As for the intensity of the southern eddy, located on the daylight side, it is, on the contrary, weaker on Ben'kova's map than on the map in Fig.73. In addition, the center of this eddy is at 12 hours 30 minutes, while on Fig.73 it is at 11 hours. Moreover, the night region of negative currents, as will be seen from Fig.77, breaks down into two distinct eddies, while in on the map of Fig.73 it reduces to a single eddy. These differences in Ben'kova's opinion should be ascribed to the difference in the initial data, with which we must reconcile ourselves, since the analysis of Fig.73 was based on the data of 21 observatories located only in middle and southern latitudes, while that of Fig.76 had only 5 observatories, but Ben'kova uses materials of 47 observatories.

The system of currents of the field S_{q2} is represented is shown for four instants of Greenwich time: 0, 6, 12, and 18 hours.

Fig.77 shows that in the equatorial latitudes at local Noon there is a region of positive current. Its development reaches its maximum (20,000 amp.) when the 300° meridian is close to the noon meridian. This region almost disappears and is displaced towards the north, when the local meridian lies on the Greenwich meridian. There is a region of negative currents on the night side. They are most distinct at 3 hours on the 300° meridian; and their maximum value in this case is 21,000 amp.

N.P. Ben'kova in her analysis also separated that the part of the field corresponding to internal causes and that part of the field (eddy) corresponding to vertical currents. But the latter problems are still controversial, since the value of the vertical currents strength obtained by Ben'kova from the analysis is many times greater than the value of the currents observed on the earth's surface.

Attempts have recently made to directly prove the existence in the ionosphere of the currents causing the diurnal variations (Bibl.48). On 17 March 1949, near the geomagnetic equator (φ = 11°S, λ =89°W) at 1120 hours and at 1720 hours local time, two rockets with magnetic instruments were sent up. These instru-

ments automatically transmitted signals by radio, at definite time intervals, of the

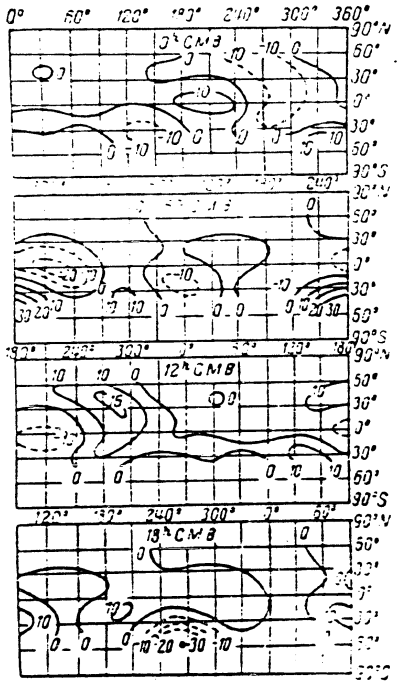

Fig.78. System of Currents for the Summer Months, According to Calculations of N. P. Ben'kova (for functions S_{q2}).

value of the magnetic field strength. The height of the rockets at these moments was determined by a radar installation on the ground and by the signals emitted from the rockets. Both rockets reached a height of over 100 km. The first rockets was sent up at a moment when the current density, according to theory (Fig.76), was maximum, and the second when the density was minimum. The results of the worked up ob-

servations in Figs.79a and 79b. The former relates to the observations at 1720 hours, the second to observations at 1120 hours. The solid lines represent the variations (decrease) of magnetic field strength with height, calculated under the assumption that the earth is uniformly magnetized and that its magnetic moment corresponds to the first term of the Gauss expansion. The dashed lines are the results

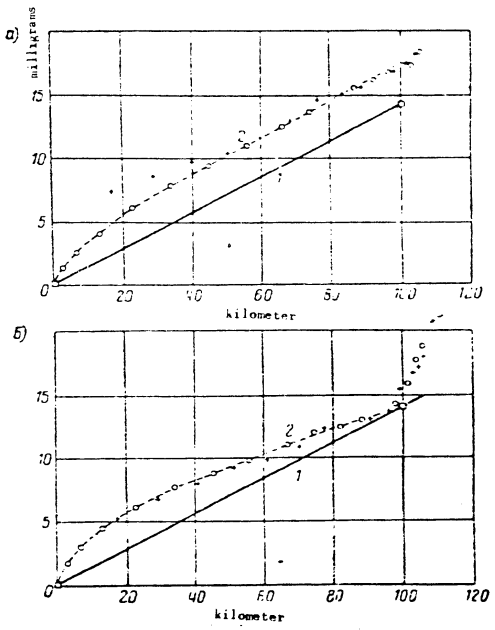

Fig.79. Variation of Magnetic Field Strength with Height from Rocket Observations:

a. at 1720 hours; b. at 1120 hours.

1. theoretical curve; 2. observed curves;

of direct observations. The points corresponding to the ascent of the rocket are marked by circles and those to the descent by crosses.

The results of these experiments indicates that currents actually do exist at a height of about 105 km, since Fig.79b has a sharp variation of the magnetic field of these heights, while Fig. 79a, correspondent to the minimum of current such variations are not observed, and the experimental curve coincides over its entire length with the theoretical.

For our final conclusions from these experiments, however, it would be necessary to perform a theoretical calculation of the variation of the field due to currents, and to compare it to the observed values, as well as to have a repeated observation of the same nature.

The noncoincidence the disagreement between the theoretical and experimental curves near the earth's surface may be explained by the existence of an anomaly in the region of the discharge of the rocket.

Section 2. Lunar-Diurnal Variations.

In addition to the variations connected with the positions of the sun with respect to the earth's surface, there also exist variations of periodic character connected with the position of Moon with respect of the horizon. The period of these variations coincides with the time interval between two successive crossings of the local meridian by the moon, i.e. with the lunar half-days.

The lunar diurnal variations are found on working up the records of magnetographs with respect to the lunar days. Since the lunar days differ from the solar days by only 50 min. 28 sec. (the lunar days equal 24 hours 50 min. 28 sec. mean solar time) it follows that to eliminate the lunar that to isolate the lunar diurnal variations there is no need to work up the magnetograms according to lunar days by taking the ordinates for each lunar hour, but that it is sufficient to use the data obtained in working them up according to solar days and a rearranging them according to lunar time.

STAT

This rearrangement consists in taking the lunar days as equal to 25 hours of solar time, and, for each hour, entering the values of the ordinates from the tables prepared for the solar-diurnal variations. Their value for the 25th hour are taken as equal to the value of the first hour of the following day. In this way each successive lunar day begins one hour later than the solar day. The beginning of the lunar day, which is established by the Astronomical Annual Yearbook is taken as the moment of the passage of the Moon through the upper meridian (the instant of upper combination). The upper lunar combination does not usually correspond to an even solar hour, but owing to the large number of days being worked up, it is sufficient, without leading to large error, to take the next solar hour as the beginning of the lunar day.

In view of the fact that the lunar day does not contain exactly 25 hours, but is 9 min. 32 sec. shorter, every sixth lunar day should have 24 instead of 25 hours; in this case the value of the ordinate for the 25th hour repeats the value for the preceding hour (cf. Appendix 2).

In addition, in order to eliminate the solar diurnal variations, up to the time of the rearrangement its mean monthly value is subtracted from each ordinate. In this way the process of "taking a mean" for the lunar day consists in eliminating all irregular variations. Since the amplitude of the lunar variations is very small by comparison to the non regular part of the variations, a considerable time interval is required to eliminate those variations. The first detailed study of the lunar-diurnal variations was made by Chapman (Bibl.49) in 1913. That author worked up by the above method the observation from the observatories at Pavlovsk, Pola, Tsi-Ka-Wei, Manila and Batavia for 7 years (1897 to 1903 inclusive).

The results of the statistical workup, and of the subsequent spherical and harmonical analysis, allowed Chapman to establish a number of regularities in the march of the lunar-diurnal variations, which differed from the regularities of the solar-diurnal.

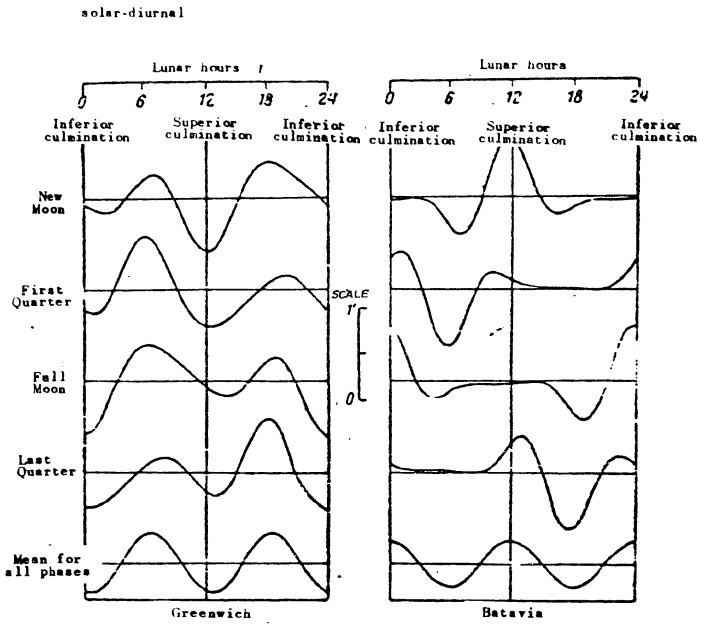

Fig.80. Lunar Diurnal Variations at Batavia and Greenwich at Different Phases of the Moon.

The principal regularity is the semidiurnal character in the changes of these variations. The curves of variation of all elements during the course of the lunar day have two maxima and two minima while the time of occurrence of the maximum and minimum vary daily during the course of the lunar months. The mean monthly curve, however, has the form of a regular double wave with maxima at 6 and 18 hours lunar time and minima at 0 and 12 hours for the northern hemisphere. Fig.80 shows the curves of the lunar diurnal variations of the declinations at Batavia and Greenwich

for the four phases of the moon, accompanied by the mean monthly curves.

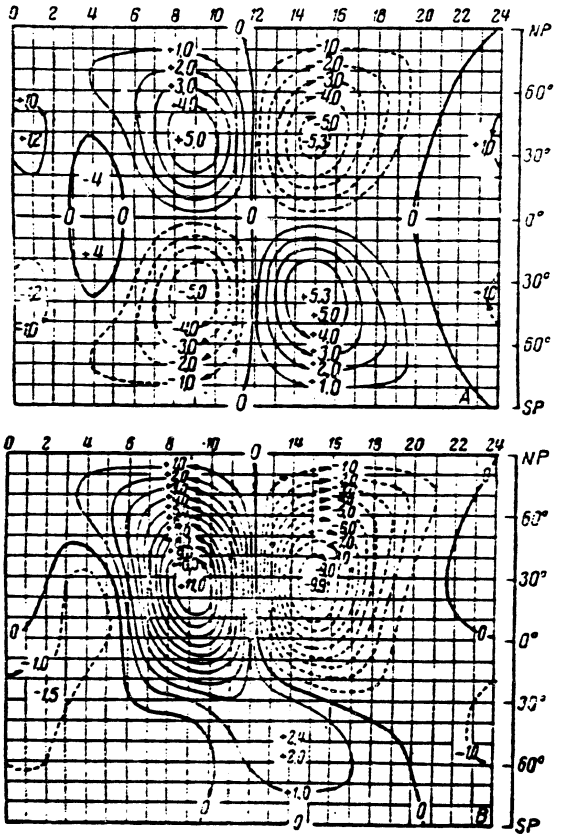

Fig.81. System of Currents Corresponding to Lunar Diurnal Variations at Equinox (upper figure) and Summer Solstice (lower figure).

The displacement of the extreme values during the course of the month may be clearly seen on the Figure. In addition, the curves for Batavia, which is located in the southern hemisphere, are almost a mirror image of the curves for Greenwich, which is in the northern hemisphere.

Chapman's expansion of the curves of the diurnal march into a harmonic series shows that the harmonics with a semidiurnal period remain constant throughout the course of the entire month, but that the harmonics of the remaining orders change their phase while the amplitude remains unchanged, thus also resulting in the displacement of the extreme values. Thus the change of phase for the first harmonics during the course of the month was -30°, for the third harmonic, $+30^{\circ}$ and for the fourth harmonic, $+50^{\circ}$.

The lunar diurnal variations have annual march depending on the position of the sun. During the time of the summer solstice the amplitude of the lunar days reach the maximum values in the northern hemisphere and the minimum values in the southern, and in the time of the winter solstice, on the contrary, they reach their minimum in the northern hemisphere and their maximum in the southern. For the vertical and horizontal components, the maximum amplitudes reach only 1 - 2 gammas, while for the declination they reach 40".

The dependence of the lunar diurnal variations on the latitude and longitude are the same character as those of the solar diurnal. Thus, with variation of the latitude, the phases remain constant, but in the northern component, on crossing the parallels $+20^{\circ}$, -20°, they change sign to the opposite. With eastern and vertical components the variation of phase also takes place on the equator. The amplitude of the observations reaches a maximum in the northern component on the equator and at latitude 45°; while the eastern and vertical components the maximum of amplitude is reached at the parallel of 20°.

The lunar diurnal variations are almost independent of the longitude. It is an interesting fact that the amplitude of the lunar variations are dependent on the

distance of the moon from the earth; more specifically the amplitude of the variations is about inversely proportional to the cube of this distance. Just as for the solar variations, spherical harmonic analysis allows the lunar diurnal variations to be explained by the existence of horizontal eddy currents, whose distribution for the new moon, is shown in Fig.81: the upper figure relates to equinox, the lower to the summer solstice. The meridians correspond to local lunar time and for the New Moon the Sun and Moon are on the meridian 12. The total current flowing in the main circuit reaches 5300 amp. for at equinox and 11,000 amp. at solstice.

Section 3. Magnetic Disturbances.

Variations without definite periods obtained as a result of subtracting the solar diurnal and lunar diurnal variations from the observed variations, and which at first glance appear entirely arbitrary, as a result of their random march, have received the name of magnetic disturbances, and at great intensity, of magnetic storms. While the amplitude of the periodic variations is expressed by a few tens of gammas, they may reach a few hundreds or thousands of gammas during the time of magnetic storms.

Magnetic disturbances may be classified according to intensity, duration, and spatial distribution, into four types.

The first type includes disturbances of very great intensity, magnetic storms occurring simultaneously over the whole earth. The amplitude of the fluctuations of the elements of the terrestrial magnetism of such storms may reach a few thousand gammas, and their duration, a few days. The second type includes disturbances of local character limited to definite regions, mainly the polar region. Local disturbances may last for one or a few hours and their intensity will exceed hundreds of gammas. While magnetic disturbances of the first type commence simultaneously over the whole earth and proceed in a single phase, disturbances of the second type, even at two nearby points, may proceed entirely different.

The third type, the bays, is the name given convexities or concavities on the

STAT

magnetograms, recalling the shape of marine bays. The bay-like disturbances, occurring simultaneously may stretch over the entire earth or may be limited to a certain region near the auroral zone.

Finally, the fourth type of periodic magnetic disturbances consists of the so-called pulsations, which are sinusoidal fluctuations in field strength with an amplitude of the order of few gammas and a period of a few minutes. The pulsations may take place simultaneously over the entire earth but may also be limited to individual regions.

In most cases, the disturbances of the first and second group usually occur simultaneously one being superimposed on the other and causing side effects, such as, for example, the induction current in the earth, which in turn yield an additional component vector of the disturbance.

Fig.82 shows magnetograms of the horizontal component during the time of the magnetic storm of 14 March 1922 according to the records of five observatories located in middle and low latitudes, while Fig.83 shows magnetograms of the horizontal component from the records of observatories in high latitudes during the magnetic storm of 19 February 1933.

The curves show that no complete parallelism is observed in the march of the elements, but that a few maxima and minima do occur simultaneously at all stations. The amplitude of the fluctuations as will be seen from Fig.82, increases with increasing latitude of the station. Thus this storm represents a storm of world-wide character with the superimposition of local disturbances.

Statistical processing of magnetic storms has allowed establishing the existence in them of at least three components differing in character and the laws of occurrence and the laws of their course.

The first of them, S_D, representing the difference $S_d - S_q$, of periodic character, with the period of a solar day, has already been considered in Section 1 of this Chapter. The second, the aperiodic variation D_{ap} is found as a result of aver-

aging a large number of world wide storms located in columns during the course of a storm, i.e. when the instant of origin of a storm is taken as the initial moment of time t = 0.

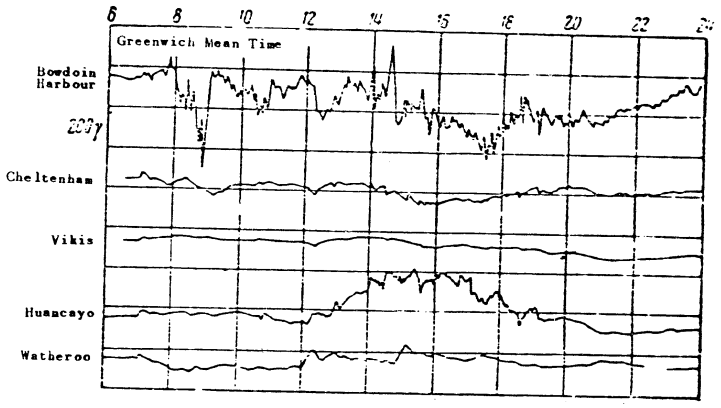

Fig.82. Magnetograms of the Horizontal Component, from Records of Various Observatories during the Magnetic Storm of 14 March 1922.

A characteristic feature of most magnetic storms is the suddenness of their appearance. Against the background of a rather quiet magnetic field, almost at one and the same instance of the entire earth all elements of terrestrial magnetism suddenly vary their values, and their subsequent course undergoes very rapid and irregular variations. For this reason it is possible to determine the beginning of a magnetic storm of the magnetograms of all observatories within 1 to 2 minutes.

The third component, which is obtained by subtracting the aperiodic variation and the disturbances of the diurnal variation from the observed ones, is really that magnetic disturbance which we term a magnetic storm and which is manifested in the

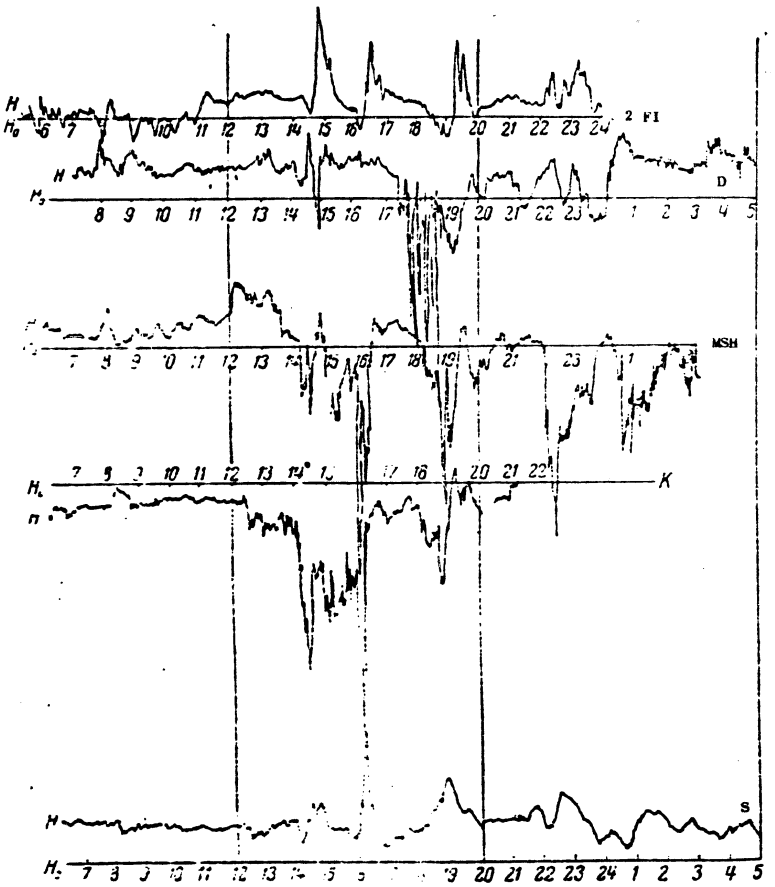

Fig.83. Magnetograms of the Horizontal Component During the Storm of 10 February 1933.

ZFI Franz Josef Land. MSH Matochkin Shar. K Kandalaksha. S Slutsk.

form of irregular rapid variations of all elements of terrestrial magnetism. This part of the variation is termed, by general agreement, the irregular component, and is denoted by D_i.

The aperiodic disturbed variation D_{ap}. A characteristic feature of this variation is that it is very distinctly manifested in variations of the horizontal component, to a lesser degree in the vertical component, and has entirely no effect on the declination.

The general character of the course of D_{ap} is as follows. The beginning of the disturbance is a short impulse which increases the horizontal component and decreases the vertical one. These variations amount to +20 gammas for the horizontal component and -5 gammas for the vertical one.

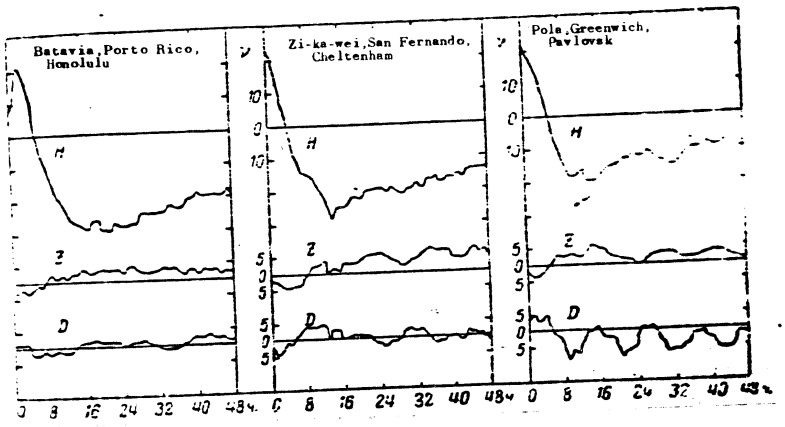

Fig.84. Aperiodic Disturbances of the Variation at Various Observatories.

The increases of values of the horizontal component lasts only a short time, from 1/2 to 2 hours. There is then a sharp fall in the horizontal component lasting

for about 6 hours and going up to 60 gammas, after which begins a long process of return to the normal state, occupying a period of time up to 2 days. This process has received the name of after disturbance.

The vertical component, after a 5 gamma drop, which likewise lasts from 1/2 to 2 hours, then begins to increase and during the course of the entire storm remains 5 γ higher than its normal value.

The declination experiences small deviations (from 1 to 2") from its normal value towards one and the same side, and therefore no regularities in its variation must be spoken of.

The next characteristic feature of D_{ap} is the dependence of these disturbances on the geomagnetic latitude of the place, and their independence of the longitude. The maximum intensity of D_{ap} is found on the magnetic equator, where the variation in the horizontal component reaches 60 gammas and more. To the north and to the south of the equator, the intensity decreases, and at latitude 60° (Pavlovsk) the variation in H amounts to 40 gammas. Further to the north, D_{ap} again begins to increase, reaching a maximum in the zone of maximum aurora.

Fig.84 shows the march of D_{ap} at three latitudes from observations of the observatories at Batavia, Porto Rico and Honolulu, located at the geomagnetic latitude 0°, Zi-Ka-Wei, San Fernando and Cheltenham, on the latitude 40°, polar, Pavlovsk, Greenwich, and Potsdam, at latitude 60°. The graph of D_{ap} shown on this Figure clearly shows all the regularities in its course that have been pointed out above.

The variations are easily traced up to latitude 60-70°. In the higher latitudes it is still not possible to isolate them so distinctly, owing to the complexity of phenomena that take place there. The local variations of high latitudes are so predominant in influence, that, superimposed on the world-wide magnetic storms, they masked the variations of D_{ap} which are of lesser intensity. This will be taken up in greater detail in the following Section.

The regularities found in the course of the D_{ap} variation, indicates that their

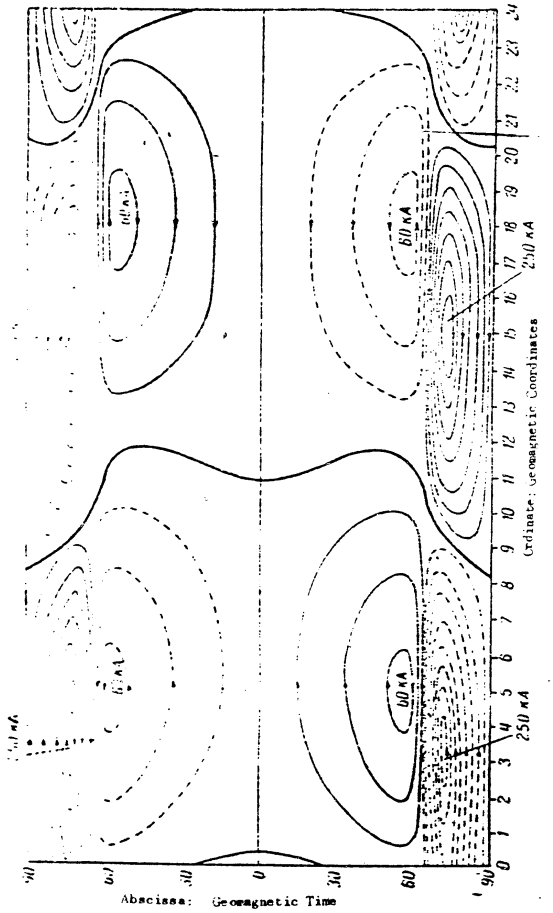

Fig.85a. System of Electric Currents Corresponding to the Variations (view from equatorial plane).

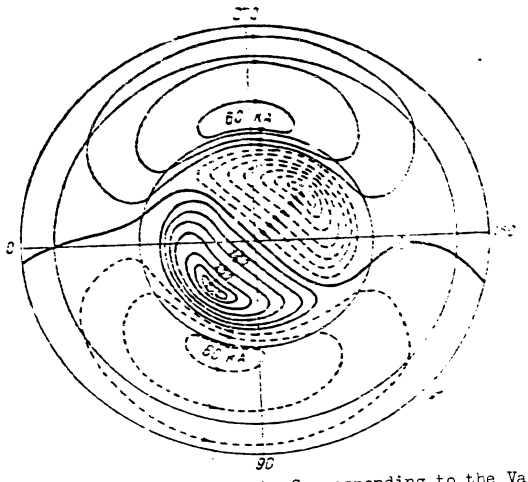

Fig.85b. System of Electrical Currents Corresponding to the Variation (view from pole).

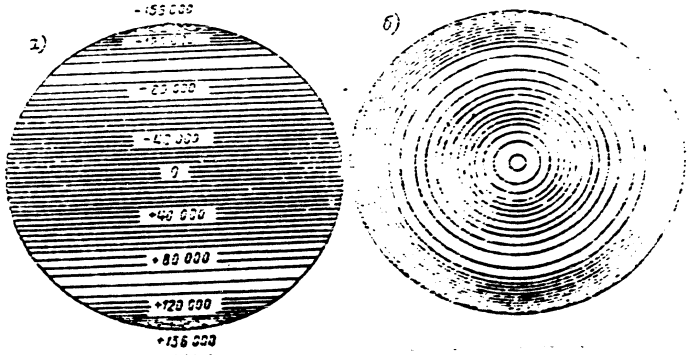

Fig.86. System of Electric Currents Corresponding to the Aperiodic Disturbance of Variation:

a. view from Equatorial Plane; b. view from Pole.

source must be the ring current flowing in the plane of the earth's magnetic equator. In this case the direction of the current in the first phase of the storm must be from east to west, and in the second phase, from west to east.

O. A. Bourdeaux was the first to make a spherical analysis of the disturbances of the diurnal variations. He succeeded in constructing a system of electrical currents in the upper layers in the atmosphere, responsible for the variation S_D. The form of this system is shown in Fig.85a and 85b, 85a represents the system of currents over the entire surface of the earth and Fig.85b a view from the geomagnetic pole. As will be seen, the maximum density of the current lines is reached in the zone of maximum auroral frequency, where the current strength reaches 200,000 amp., while the total current over an area equal to one quarter of the earth's surface does not exceed 40,000 amp. The strong crowding of the current lines is also observed in the polar regions between the auroral zone and the pole where the total current amounts to 270,000 amp. Fig.85a showed the system of current consists of eight closed currents of which four are located in the eastern hemisphere and four in the western. The system of currents constructed by Bourdeaux is in rather good agreement with the observed diurnal march of the variations. Thus, for example, it reproduces the inversion of the X component of the latitude 55°, the hours of maxima and minima of the X and Y components, and the general march of the variations in high latitudes. There is still no theory of the origin of this system of currents.

The system of currents corresponding to the aperiodic disturbed variation, viewed from the sun and viewed from the pole, is shown in Fig.86, from which it is clear that the system forms a current flowing along the surface of a sphere parallel to the equator, its current density declining from the equator to the auroral zone, and increasing on passage through the auroral zone, reaching a maximum between the pole and the zone. The theory of the aperiodic variation has been developed by Chapman and Ferrari; and it will be set forth in the following Section.

Bay shaped variations. As stated above, the bay shaped variations which we

shall denote by D_B, have the shape of the shore line of marine bay on the magnetogram records, with an amplitude reaching a few hundred gammas. The bay-shaped disturbances appear most distinctly in the horizontal component. In this case D_B may arise as a solitary disturbance amidst a quiet field, or may be superimposed one on the other, or may also be present in a general magnetic disturbance. Fig.87 shows the character of a record of such disturbances observed at the observatory at Kew, England, in February 1911. Owing to the existence of individual bay-shaped disturbances,

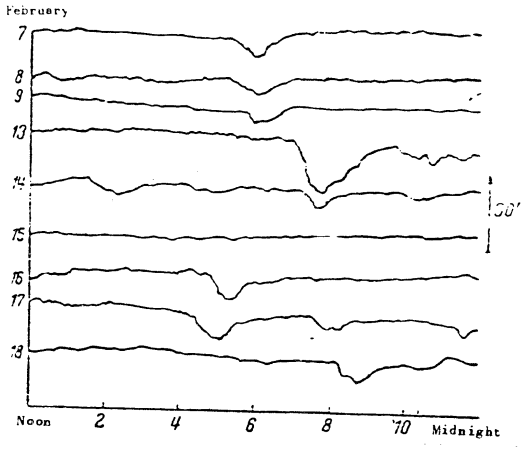

Fig.87. Bay-Shaped Variations at the Western Observatory at Kew, (England).

undistorted by other variations, in the records of the observatory, their more detailed study was possible. The observations show that D_B, occurring simultaneously at all stations of the world, have a maximum intensity in the auroral zone where the amplitude of the variations of H is tens of times as great as in the low altitudes.

If a graph of the variations of H and Z during bay disturbances is plotted against the latitude, laying off the maximum amplitude of D_B along the axis of ordi-

nates and the latitude of the stations located close to one and the same meridian are laid off on the axis of abscissas, it will take the form shown in Fig.88. This graph is analogous in its shape to the graph of the magnetic field excited by a magnet with its axis parallel to the meridian and its center at latitude 70°.

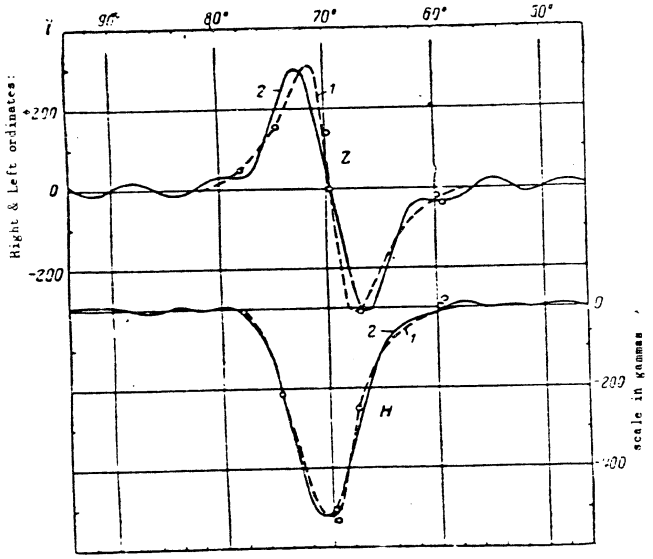

Fig. 88.

1. observed curve passed through the points; 2. approximation for observed curve by the aid of six times of a trigonometric expansion.

In view of the fact that such graphs are followed through simultaneously on other meridians as well, it is necessary, for the formation of D_B to postulate the existence not of a single magnet but of a series of magnets parallel to each other

and located along the parallels, so, in the alternative, the existence of an electric current equivalent to it and flowing in a narrow beam which may be taken as a linear current along the 70th parallel.

Indeed, a linear current flowing at a distance R from the earth, creates on the earth's surface in a direction perpendicular to the current, a magnetic field determined by the Biot-Sawara law.

The components in this field along the vertical and horizontal will obviously be:

$$\delta H = \frac{2IR}{R^2 + x^2}, \qquad \delta Z = \frac{2Ix}{R^2 + x^2}.$$

These formulas coincide with equation 8.45 for a single-pole filament (page 349), the graph at which are given in Fig.167. By comparing these graphs with the curves without the D_B variations in Fig.88, and obtained from observations, it will be seen that they are in close agreement with each other.

If it is postulated that D_B is caused exclusively by current flowing in the upper layers of the atmosphere, then the height at which it flows, and the magnitude of the current strength, can both be easily determined.

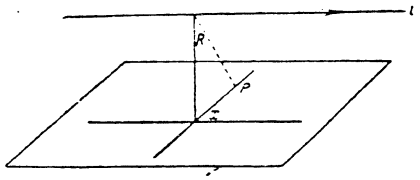

Fig. 89.

The height h of the linear current is found graphically by constructing the field strength vectors at points located along the meridian near the 70th parallel, and producing perpendicular to them. The point of intersection of the perpendiculars

will give us the linear current on the scale of the height, since for the linear current the direction of the magnetic field is perpendicular to the distance between the linear conductor and the given point of the field.

The current strength I is determined from the Biot-Sawara law for a linear current:

$$\delta H_T = \frac{0.2I}{h},$$

where δH_T is the field strength of D_B at the 70th parallel.

Such determinations gave a value of the order of a million amperes for the current strength and a value ranging from 100km to a few hundred kilometers for the height.

Further refinement showed that the sources of the D_B variations do not consist only of the currents in the atmosphere, but also of induced currents within the earth, which is responsible for about 40% of the entire field of variations.

Magnetic pulsations. Pulsations represent regular fluctuations of the elements of terrestrial magnetism, mainly of the declinations and the horizontal component, with a period from 20 sec. to a few minutes and an amplitude of a few gammas. At the usual rate of rotation of the drum of 20 mm per hour, they are found on magnetograms in the form of a sawtooth curve with small peaks. At a higher rate of rotation of the drum, however (3 mm per minute) they are recorded in the form of a regular sinusoid. A characteristic feature of these variations is that they are observed with mainly around midnight, often from 22 to 2 hours of the following day. The usual amplitude of these variations does not exceed a few gammas, although pulsations with larger amplitudes are sometimes observed. Thus, for example, on 12 September 1930 at the observatories of Abisko (Lat. = 68.4°, Long. = 18.8°), and Trömso (Lat. = 69.7°, Long. = 18.9°) in Norway, pulsations with amplitude as high as 30 gammas were registered. Fig.90 gives the records of these observatories showing the character of the pulsations themselves. Since pulsations are of a sinusoidal form, it has been postu-

lated that they are not the thought has been expressed that they might perhaps be natural oscillations of the variometer magnets themselves (cf. Chapter XII), due to the seismic vibrations of the soil. But pulsations are also observed with magnets with periods of natural oscillations far from those of seismic vibrations.

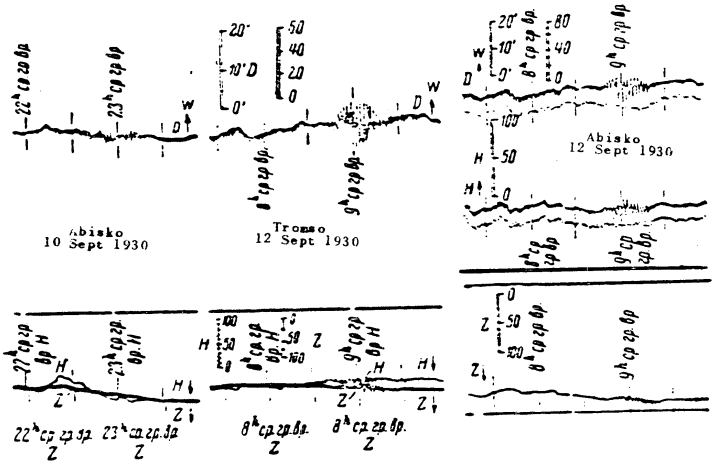

Fig.90. Magnetic Pulsations at Abisko and Tromso, 12 September 1930.

An important factor is also that the pulsations are sometimes observed simultaneously at several stations, such as, for example, the pulsations recorded at Abisko and Tromso and shown in Fig.90, which are also observed at a number of other stations. This has given ground for postulating the world-wide character of the causes responsible for pulsations. A more detailed study of the special distribution of pulsations, however, showed that the region which the action of these causes extends is limited to a radius of not more than 1000 km. Even such a gigantic pulsation as the one observed on 12 September 1930 at Tromso, which is at a distance of not more than 100 km from Abisko, declines by a factor of several kinds by comparison with Abisko.

A proof that the pulsations constitute a real phenomenon of nature is also their record by instruments based on the conducti n principle, i.e. instruments which react to a variation of the magnetic field with time.

In spite of the large number of works devoted to the study of pulsation, their causes have still not been established and there is no theoretical explanation for them.

Section 4. Variations at High Latitudes.

The systematic study of the variations in high latitudes began somewhat more than 20 years ago, when, in connection with the Second International Polar Year (1922-1933) a number of magnetic observatories were open in these latitudes. The extraordinary complexity of the variational phenomena in these parts of the earth, however, demands for its study a large observational material, both in time in a number of observatories. For this reason the relatively short period of observations and a small number of observatories still fail to make it possible to establish definite regularity in the course of the magnetic variations at these latitudes, as for the low and middle latitudes, but even the materials that is available at the present time still allows us to point out certain features and to find regularities of one kind or another in these phenomenon. The most complete material for the past 25 years has been collected by the Arctic observatories of the USSR. This material enabled A. P. Nikol'skiy to find new phenomena in the diurnal march of magnetic activity in high latitudes. Very valuable material was secured during the Second International Polar Year, when a whole system of temporary station operated on a single common program. The results of the observations of this sphere were worked up by E. Vestine and published in 1947, in the form of many different graphs with explanatory text and a few conclusions.

The results of this workup are the materials which must serve as the foundation for future studies and deductions. But certain conclusions may already be made, even on the basis of this material.

STAT

The solar-diurnal variations S_q. One peculiarity of the diurnal variations in high latitudes is their dependence not on the geographical latitude, but on the magnetic. This dependence is manifested with particular sharpness on passage through

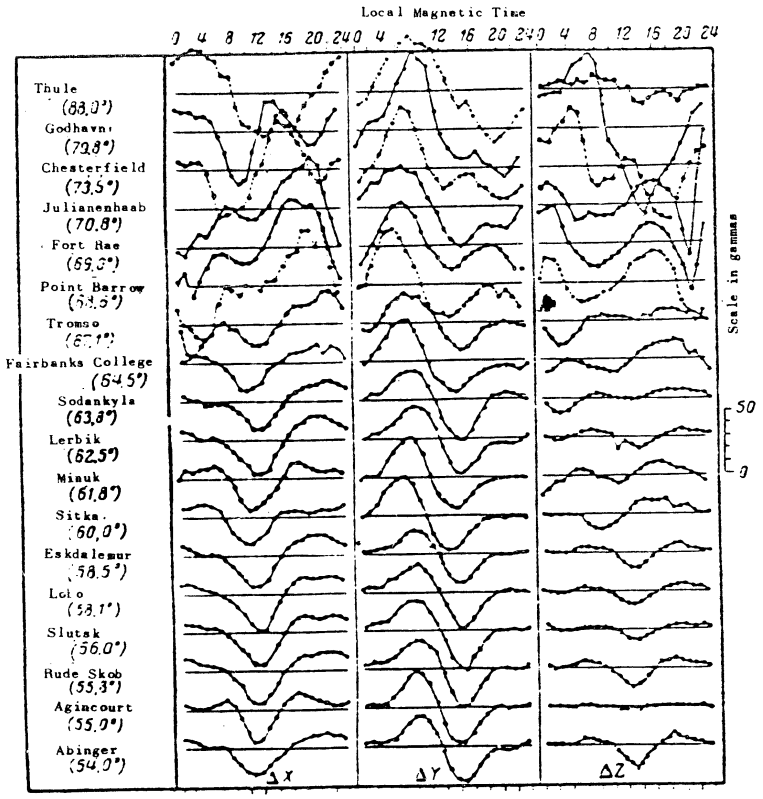

Fig.9la. Diurnal March of the Elements X, Y and Z on Quiet Days During the Period of the Summer Solstice at High Latitudes.

the zone of maximum aurora frequency. A second peculiarity consists in the considerably higher value of the amplitude and phase variation in the vertical and northern

component on passage across certain magnetic parallels.

Fig.91 shows the mean diurnal course march of the variations on quiet days in the winter and summer months, as well as the mean annual values at the stations of

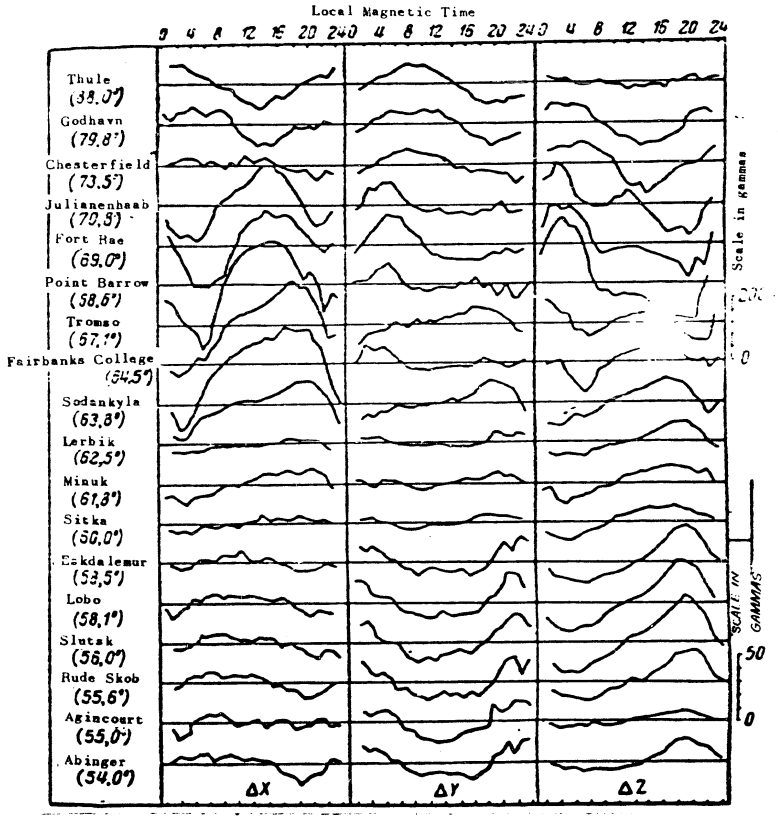

Fig.91b. Diurnal March of the Elements X, Y and Z on Quiet Days During the Period of the Winter Solstice at High Latitudes.

the Second International Polar Year, located in the high latitudes, and at the same

time at a few permanently operated observatories in the middle latitudes. These curves clearly indicate the latitudinal dependence of the variations, as well as their dependence on the season. In the winter months the variations are considerably

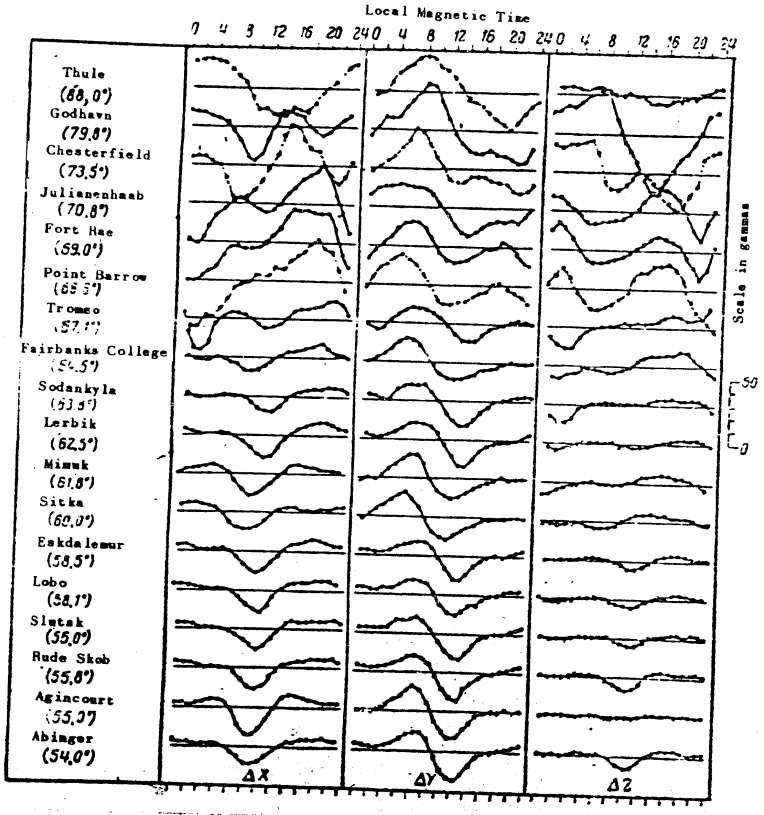

Fig.91c. Mean Annual Diurnal March of the Elements X, Y and Z on Quiet Days in High Latitudes.

less than in the summer. In addition, our attention is struck by absence of that

regularity in the march of the variations which is observed in the middle and low latitudes, which phenomena is probably due to the superimposition if irregular dis-

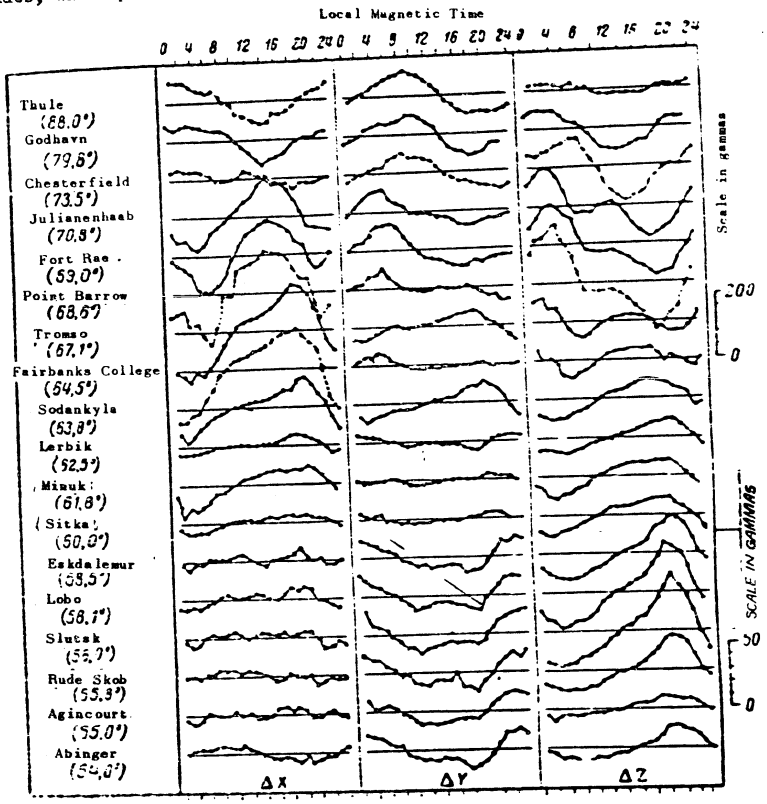

Fig.92. Diurnal Variations on Stormy Days S_D, in High Latitudes (Mean Values for the Year).

turbances, since in the high latitudes, disturbances are observed even on the quiet-

est International days.

As a result of this, the variations S_D are of extremely great interest, i.e.

the diurnal variations on disturbed days minus, those on the quiet days.

Fig.92, showing the mean annual march of the S_D variations, shows that the S_D variations are of more regular character than the S_q. Their peculiarity is the change in phase of the Z component in passing through the zone of maximum aurora frequency, located approximately at the magnetic parallel 68°, and also the maximum value of the amplitude of the X component in this zone. The change of phase in the X component takes place at magnetic latitudes 51°, and then again at latitudes 72°. The eastern component likewise changes its phase twice, once at latitudes of about 60° and afterwards in the auroral zone. The amplitude of all the components reaches values in the auroral zone that many times exceed the values in the middle latitudes.

While the variations S_q have a strong seasonal dependence, the variations S_D remain almost the same throughout the entire year.

The aperiodically disturbed variation. The least studied variation in the high latitudes is the aperiodic variation (D_{ap}), as a result of which there is still no well established regularity in its course. But the result of the workup of 11 storms registered at the Arctic stations during the second polar year allow us to give a few conclusions. In view of the small number of storms, they have been worked up for a group of stations located close to one and the same magnetic latitude. The result of this workup is given by Fig.93 in the form of graphs, which show that the northern component at the beginning of the storm has a value somewhat lower than normal, and which then during the first 10-12 hours, falls to a minimum, after which it begins gradually to increase to its normal level. The absolute value of the minimum increases on approaching the auroral zone, where it reaches 80 γ. Inside this zone it declines, and at latitude 84° it reaches a value of 30 γ, in this case the character of the curve changes completely and takes on the form of a periodic curve.

The eastern component varies periodically, now increasing now decreasing, within the range of 10-20 γ, indicating the absence of any definite regularities, and in all probability, also indicating the absence of components of D_{ap} in this di-

rection, since the fluctuations might be caused by the superimposition of random variations.

The vertical component, up to the auroral zone, has the same character and the same value of the amplitude as in the middle latitudes. On passing through the auroral zone, the vertical component, 10 hours after the onset of a storm, begins to increase sharply, reaching a maximum of 60 to 100 gammas within a few hours. The highest value of the maximum is assumed at latitude 71°.

Magnetic activity or degree of disturbance. The existence of a large number of magnetic disturbances in the Arctic latitudes, which, being superimposed on the periodic variations, make it impossible to isolate them by an averaging, force us to seek other methods of studying these disturbances than the study of the diurnal variations. Such a method, proposed by A. P. Nikol'skiy (Bibl.50) in 1938, consists in taking the mean values of the hourly values of some measure of activity, either the international characteristic k, or the characteristic proposed by Nikol'skiy himself, the length of the curve D. The mean value of this activity over a certain time interval was termed the degree of disturbance by Nikol'skiy, who studied it in its relationship to the time, latitude, and longitude of the place, season, and the 11-year cycle of solar activity. These studies yielded new material for the elucidation of the causes of magnetic disturbances, and established a number of new regularities in the course of these variations.

The principal result of Nikol'skiy's work was the establishment of a definite regularity in the diurnal march of the magnetic of the degree of magnetic disturbance, which was as follows: the degree of magnetic disturbance during the course of the 24 hour day has two maxima, one of which comes in the morning hours and occurs according to universal time, the other in the evening hours, and occurs according to local time.

In addition, the instant of onset of the morning maximum in the eastern hemisphere depends linearly on the geomagnetic latitude of the station, while in the

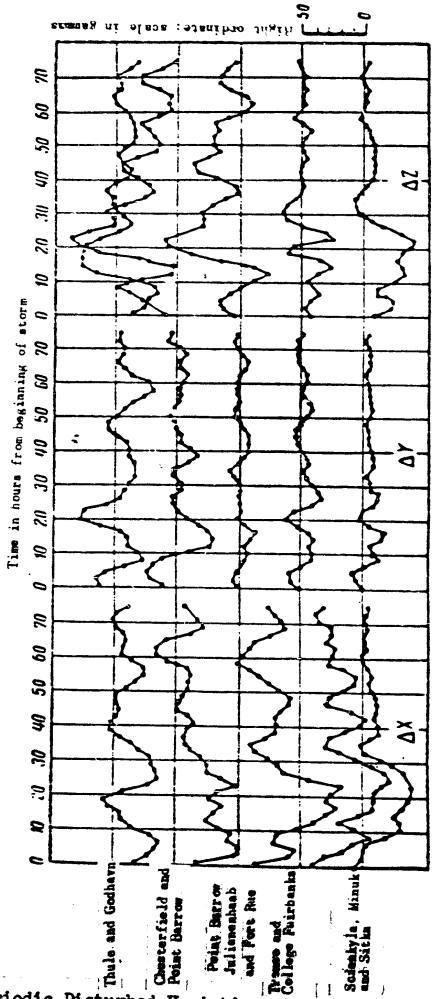

Fig. 93. Aperiodic Disturbed Variations in High Latitudes.

western hemisphere this instant of onset takes at one and the same hour (1530-1630 hours Greenwich mean time).

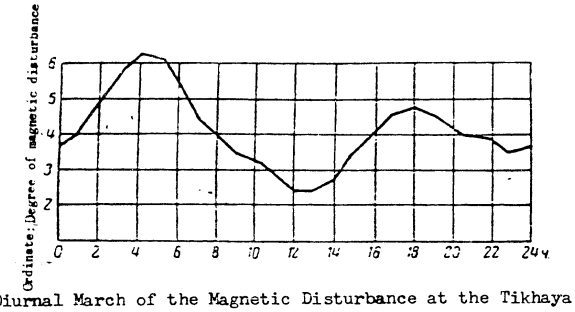

Fig.94. Diurnal March of the Magnetic Disturbance at the Tikhaya Bay Observatory.

Fig.94 shows a typical curve of the diurnal march of disturbance at the Tikhaya Bay Observatory (Lat. 71.5°, Long.= 153.3°) while Fig.95 shows the relation of the

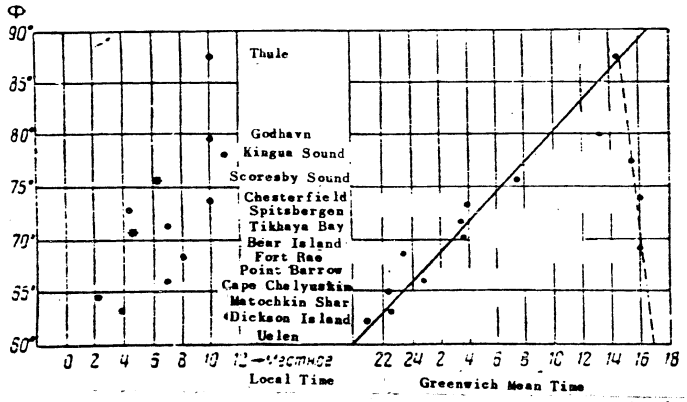

Fig.95. Relation of the Morning Maximum of Disturbance to the Universal Time (according to Nikol'skiy).

morning maximum to universal time (Greenwich mean time). These curves thus give us

grounds for asserting the correctness of this regularity.

On investigating the behavior of these maxima in relation to various parameters, Nikol'skiy found that each of them behaves differently, which gave him reason to enunciate the hypothesis that these maxima were of different nature.

Thus, for example, the value of the morning maximum is almost independent of the season, while the evening maximum has its lowest value in winter and its highest value in summer, which is illustrated by the curves of the diurnal march of disturbance at Tikhaya Bay, worked up by seasons — winter, summer, and equinoxes (Fig.96).

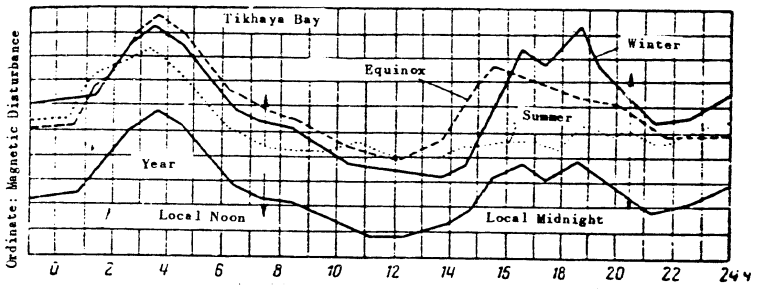

Fig.96. Relation of the Diurnal March of Magnetic Disturbance and the Season.

The value of both maxima of disturbance depends further on the general state of the magnetic field.

With the increase of the general degree of disturbance, the value of the maximum increases, but in different ways. Fig.97 shows the diurnal march at Tikhaya Bay for six groups of days with varying degrees of disturbance. It will be seen that at the beginning, when the total disturbance is small, the morning maximum is predominant, but later, with the increase in the total disturbance, both maxima increased, but the night one increases considerably faster, and on stormy days is almost twice as large as the morning maximum.

Finally, we point still another peculiarity, the dependence of the value of the

maxima on the geomagnetic latitude. This dependence is shown in Fig. 98, which indicates that with increasing latitude the maxima also increases, but that the evening maximum reaches its highest value in the auroral zone while the morning maximum reaches it at the geomagnetic pole.

All these facts confirm beyond the doubt the thought that these two maxima are

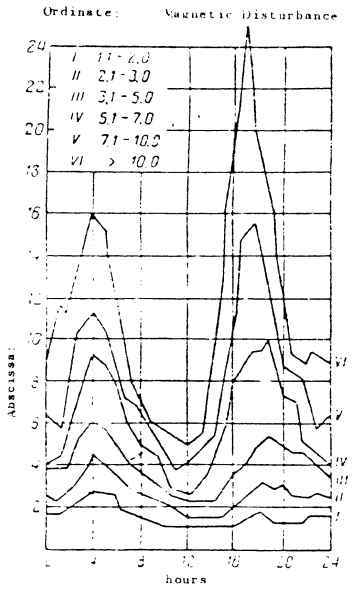

Fig. 97. Relation of the Diurnal March of the Magnetic Disturbance on the Total Disturbance of the Field.

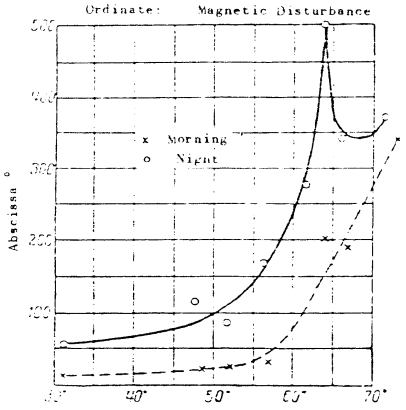

Fig. 98. Relation of a Maximum of Disturbance to the Geomagnetic Latitude.

due to different causes. However, attempting to find these causes and to explain the diurnal march of the degree of the magnetic disturbance, Nikol'skiy constructs a very simplified scheme of currents postulating that the variations are due to a rectilinear current of corpuscles of different signs, with the morning maximum attributed to a stream of corpuscles of one sign, while the evening maximum is attributed to a steam with the opposite sign, although he adduces no quantitative calculations

to confirm this hypothesis.

Nikol'skiy further attempts to cast doubt on the methods of segregating the diurnal and aperiodic disturbed variations, and claims in his work that, at least in the high latitudes, the diurnal variations and aperiodic disturbed variations do not exist as actual processes and that they represent fictitious phenomena obtained as a result of the statistical treatment. It is premature to agree with such theories and conclusions, since they are constructed with no quantitative analysis whatever and are unconfirmed even by elementary mathematical calculations.

Section 5. The Aurora.

The magnetic disturbances observed on the earth, and especially in the Arctic regions, are closely connected with the aurora. This connection is not merely external, but is also internal and physical and is due to the common causation of both phenomena. For this reason, in considering the causes of magnetic disturbances they cannot in any case be abstracted from the fact of the aurora, which allows us to understand more clearly and more profoundly the nature of the magnetic disturbances.

1. The forms of the aurora. The aurora may be classified according to its form into two great groups: auroras of nonradial structure and auroras of radial structure. Each of these groups in turn subdivided into a number of subgroups.

Those of nonradial structure include the following:

1. uniform quiet arcs the lower edge of which is sharply pronounced, while the upper edge is blurred. A dark segment is observed between the lower edge of the arc and the horizon;

2. uniform bands usually extending in the same direction as the arcs, but less regular in form. The lower edge is often sharply defined, but is irregular in form. In most cases the bands are broken up by dark spaces and for this reason have the form of feathery clouds;

3. pulsating arcs, which rhythmically appear and disappear with a period of a few seconds;

4. diffusely luminous surfaces having the form of a veil covering a large part of the sky;

5. pulsating surfaces consisting of diffusely luminous parts of the sky, appearing at one and the same place.

The radial-structure auroras include the following:

1. arcs with a radial structure;

2. bands of radial structure, seemingly uniform, but consisting of a number of streamers;

3. drapery, consisting of a few bands with very long streamers, similar to a folded curtain. The lower edge of the drapery is usually more illuminated. Near the magnetic zenith (the projection of the magnetic pole on to the celestial vault; it has the form of a fane;

4. isolated streamers which may be narrow or wide, short or long, isolated or in the form of beams;

5. a Corona, consisting of streamers, bands, or drapery, converging to a single point near the magnetic zenith.

2. The direction of the streamers of the aurora. A substantial factor in the observations of the radial structure is a sufficiently close coincidence of the direction of the streamers with the direction of the lines of force of the earth's magnetic field. This fact is one of the basic facts for the construction of the modern theory of the aurora, based on the motion of charged particles in the earth's magnetic fields.

Indeed, by observing visually the streamers of the aurora it may be noted that all of them converge on a single point located near the magnetic zenith. The results of precise determinations of the direction of the streamers by photographing the Corona of the aurora, are given in Table 19 which gives the observed height and the azimuth of the point of convergence of the streamers and of the magnetic zenith.

As will be clear, the coincidence between the point of convergence of the

STAT

streamers and the magnetic zenith is observed within the limits of accuracy of the observations themselves.

Table 19

Station	Observer	Year	Number of observations	Point of convergence		Magnetic zenith	
				h	a	n	a
Holde	Vegard and Krogness	1914	11	75.4°	−2.7°	76.7°	−2.5°
Oslo	Stoermer	1917-1921	9	70.0	−9.8	70.8	−9.7

Vegard, studying the structure and distribution of the light along a bundle of rays, came to the conclusion that the rays always follow the direction of the lines of force and that the point of their intersection corresponds to the magnetic pole, which may be displaced owing to the appearance, during the time of strong auroras, of an additional magnetic field formed by the ring currents around the earth.

3. The height of the aurora. The height and position of the aurora in space may be determined by the simultaneous observation of their coordinates (altitude and azimuths) at two points, the distance between which is known.

The first determinations for this method were made over 200 years ago and they have since been repeated by many investigators, who have given the height within the range of 80 to 200 km. But exact determinations of the height of the aurora became possible with the introduction of the photographic method.

This method was first used by Störmer (Bibl.51) in 1910, and in 1913 at the Observatory at Vessekop in Lapland, and is as follows at two stations, working simultaneously and having a telephone connection between them, the aurora is photographed on motion-picture film. Stars are photographed together with the aurora on this film. By determining the position of the aurora among the stars on the film, and knowing the distance between the stations and the azimuth of the aurora, its position in space can be determined with fair accuracy and its height can be calculated. The

length of the base, or distance between the stations, was only 4.5 km in 1910, but in 1913 it was as much as 27.5 km, which made it possible to make more exact determinations.

Thanks to the short exposure (half a second or less), this method allowed the observation not only of quiet auroras but also of rapidly pulsating ones.

The result of the first determinations in 1910 - 1913 show that the height of

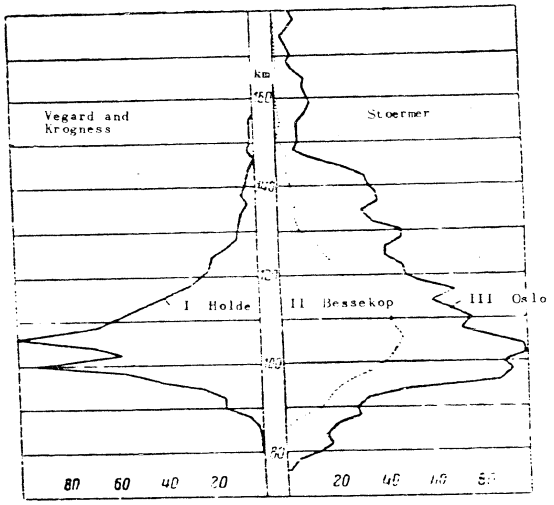

Fig.99. Probability of Appearance of Lower Boundary of Aurora Plotted against Height.

the aurora ranges from 87 to 350 km, with two maximum frequencies being observed, one at height 101 and the other at height 106 km.

These results were confirmed in 1920 at the Holder observatory by more systematic determinations of the upper and lower boundaries of various types of aurora.

The results of these determinations are shown in Fig. 99 in the form of curves showing the probability of the appearance of the lower boundary of an aurora at one height or another. These curves show clearly that the most frequent lower boundary of the aurora is at height 100-106 km.

During the period from 1910 down to the present, thousands of determinations of the height and positions of the aurora in space have been made. This material has allowed the following conclusions to be drawn. The quiet forms of the aurora, diffuse arcs, bands, and pulsating areas have on the average their lower boundary at a lower height than the pronounced radial structures. Moreover, the mean height of the lower boundary of the quiet forms and drapery is almost independent of the latitude of the place, but for the forms of radial structure the mean height increases from the auroral zone (cf. infra) towards lower latitudes. The results of measurements of the lower boundary of various forms of the aurora at various observatories is given by Table 20.

Table 20

Type of aurora	Lower Boundary of the Aurora					
	Holder, Lat 69°56', Long 62°55'		Tromsø (Lat 69°49', Long 16°57')		Oslo (Lat 60°0', Long 10°40')	
	Height, km	Number of observations	Height km	Number of observations	Height km	Number of observations
Streamers	113.2	61	117.0	127	146.9	119
Drapery	103.8	409	112.9	1039		
Drapery-like arcs	106.6	888	106.7	1175	100.0	150
Pulsating areas		100	107.3	66	—	—
Diffuse arcs			—	—	118.5	201

The upper boundary of various types of the aurora lies within the limits of 140 to 250 km. The auroras in the form of streamers are the longest. Thus the mean

height of the upper boundary of some forms are as follows: streamers, 250 km; drapery, 176.3km; drapery-like arcs, 174.4 km; diffuse arcs 143.4 km.

Thus the length in a vertical direction is 137 km for the streamers, 37 km for the drapery, and 34 km for the diffuse arcs.

But auroras are also observed in the form of streamers whose lower boundary goes up to a few hundred kilometers and its upper to a thousand.

Thus Störmer registered on a September 1926 photogram at Oslo an aurora in which the lower boundary lay from 200 to 400 km, and the upper boundary from 1000 to 1100 km.

4. The geographic distribution of the aurora. As indicated above, the aurora is observed not only in the Arctic regions but also in the middle latitude and even lower. Thus, for example, in the 1870's an aurora extended to Egypt and even to India. But the frequency of appearance of the aurora (the number of displays per year) is very small in the middle and low latitudes, while in the high latitudes the aurora is observed almost daily.

As early as 1881 (Bibl.52) a map of the isolines of auroral frequency (isochasms) was prepared and showed that the aurora was most frequent in a region about 23° distant from the geomagnetic pole. This zone has received the name of zone of maximum aurora.

For 60 years from the publication of this map an immense material was accumulated on observations of the aurora, which made it possible in 1947 (Bibl.53) to construct a new map of the isochasms (Fig.100) which on the whole repeats the 1881 map. A similar zone of maximum aurora exists also around the south geomagnetic pole.

On the map, the isochasm corresponding to the maximum frequency is shown by a thicker line. The values of the isochasms are shown in arbitrary units, in percent related to the maximum zone.

Fig.101 is a map of the earth on which the northern and lower zones of magnetic of maximum auroral frequency and the geomagnetic equator corresponding to the uni-

form magnetization of the earth, are shown.

The following we shall now indicate a few of the most of the strongest auroras

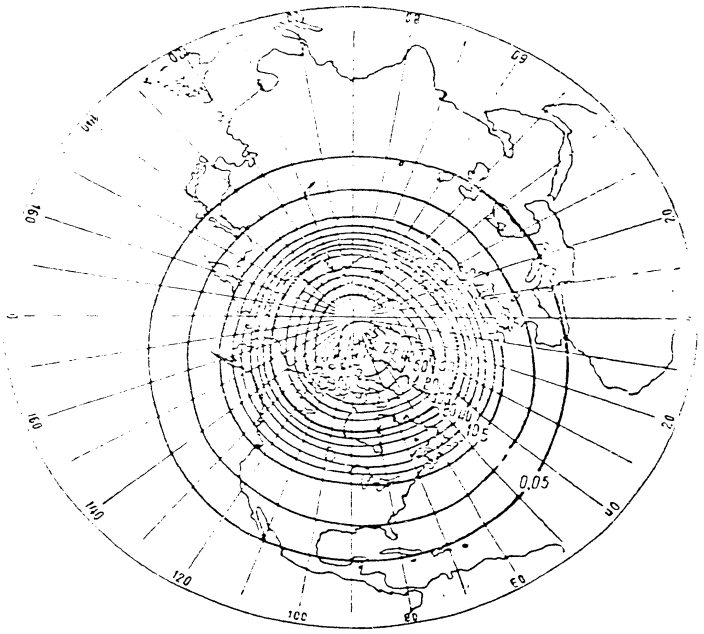

Fig.100. Isochasms -- Isolines of Auroral Frequency.

during the last 100 years. The strongest of them was on 4 February 1872 when it was visible at Bombay, Lat. 19° N, at a distance of 80° from the magnetic pole. The magnetic zenith of this aurora was observed at Constantinople and Athens. The aurora australis was observed simultaneously at Lat. 20°S, at a distance of 72° from the earth's magnetic pole. The next intense aurora was observed on 14-15 May 1921 when the aurora australis reached the Islands of Samoa (13.8°S). This aurora was accompanied by strong magnetic storms.

STAT

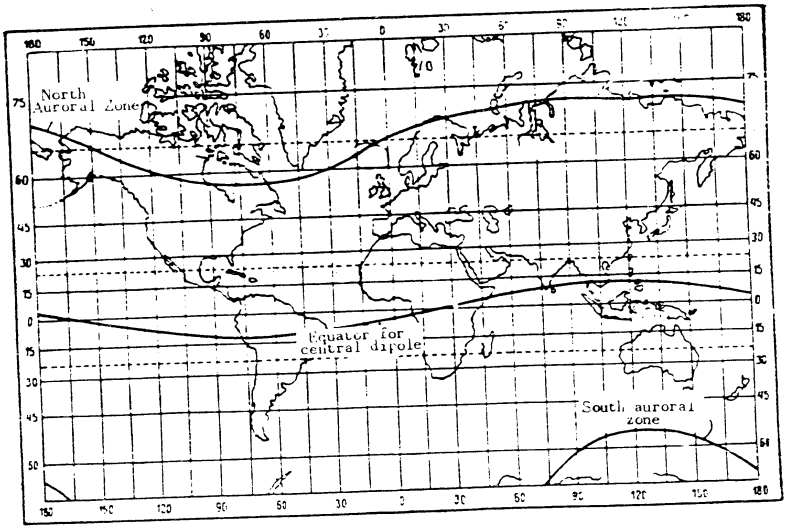

Fig.101. Zone of Maximum Auroral Frequency.

STAT

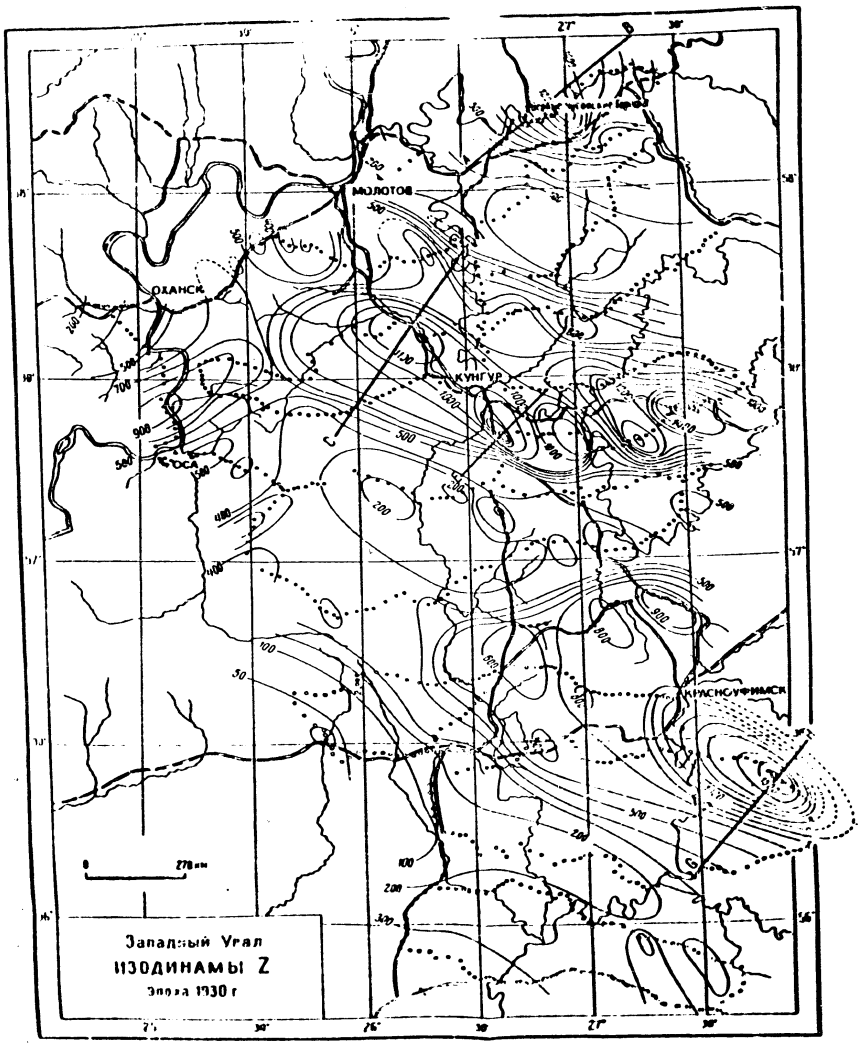

Fig.48.

Strong auroras were also observed in January and April, 1938.

5. The diurnal distribution of the aurora. By observing the aurora daily it may be noted that during the 24 hour day it appeared not at random but have a tendency to group themselves about a certain moment about a certain time. Thus, the observations during the First International Polar Year, 1882-1883, showed that auroras in the form of streamers draperies and corona have a distinct maximum in the evening hours, and a winter maximum in the morning hours. In this case, the moments of maximum appearance of the aurora, according to local time differ at the various stations, depending on the geographical coordinates. Thus at Bossekop station (Lat. 69.57° and Long. 23.15°E) this maximum was at 2125 hours local time, while at Fort Rae, Lat. 62.39°, Long. 115°49'W it was at 2400 hours.

If, however, the local magnetic time is taken instead of the local solar time, then the times at maximum auroral frequency will be the same for all stations, and correspond to 23 hours magnetic time. In this case, "by magnetic time" we mean the angle between a plane passing through the magnetic axis through the earth and the sun and the plane passing through the same axis and the given station.

6. The spectrum of the aurora. The spectroscopy of the aurora, owing to its low intensity, requires special spectrographs with a large f number and an exposure measured in tens of hours. For this reason it was long impossible to identify some of the lines observed with the lines of any specific element. In 1912 Vegard succeeded in obtaining 33 lines by the aid of such a spectrograph and an exposure lasting a month.

Subsequent studies showed the presence of a large number of lines in the auroral spectrum in the visible, infrared and ultraviolet regions. Among these lines, the most intense is the green line at 5557.3 A. This line does not correspond to any of those observed under laboratory conditions.

Since this line is observed not only in the spectrum of the aurora but also in the spectrum of the luminescence of the night sky in the absence of the aurora, it

gave Vegard cause for enunciating the hypothesis that solid crystaline particles of nitrogen were present in the upper layers in the atmosphere and on bombardment by electrons luminesced and radiated this line.

But this green line at 5777 A (sic) was later obtained under laboratory conditions from atomic oxygen, and in this way the problem of the origin in this line of the spectra of the aurora and the night sky was solved.

Of the other brightest lines, we may note the lines at 3914, 4278, 4708 and 5225 A, belonging to the ionized nitrogen molecule N_2^+ and the lines 3997 and 4059 A, belonging to the neutral nitrogen molecules.

The discovery of the green line in the spectrum of atomic oxygen permitted the prediction of three red oxygen lines at 6300, 7364 and 6392 A, of which two were found in the auroral spectrum.

Thus the aurora on the whole is the luminescence of atomic oxygen and molecular nitrogen of which the ionosphere is composed.

7. The connection between the aurora and solar activity. The connection between the magnetic disturbances and the aurora was first established as far back as the beginning of the 18th Century, when it was noted that the aurora are accompanied by magnetic storms.

Further observations completely confirm this discovery, but it has still not been possible to establish the essential functional connection between these phenomena. All the studies in this direction lead merely to a statistical correlation which always was found to be high. Thus the regular observations at the magnetic observatory at Tikhaya Bay made in 1932-1933 allowed a Soviet investigator (Bibl.54) to compare the magnetic characteristics 0 1 2 3 4 with the auroral characteristics 1 2 3 4. In the auroral characteristics, Unit one corresponds to the absence of an aurora on that day, two to the presence of moderate aurora, but without a radial structure, three, to the presence of bright auroras of radial structures, and four to very bright auroras. The results of these comparisons are given in Table 21, which

indicates the number of days with the corresponding auroral and magnetic characteristics.

Table 21.

Comparison of Magnetic Activity and Auroral Activity from Observations at Tikhaya Bay During the Period from October 1932 and March 1933.

Intensity of aurora (characteristics)	Total number of cases	Number of cases when the magnetic characteristics was					Mean magnetic characteristics
		0	1	2	3	4	
Faint (1)	186	10	114	46	7	3	1.3
Moderate (2)	76	3	27	32	9	5	1.8
Bright (3)	31	0	11	10	5	5	2.1
Very bright (4)	10	0	3	2	2	3	2.5

These results, in general, confirm the absence of magnetic disturbances on those days when aurora are not observed and their appearance with the appearance of the aurora, although there are also exceptions.

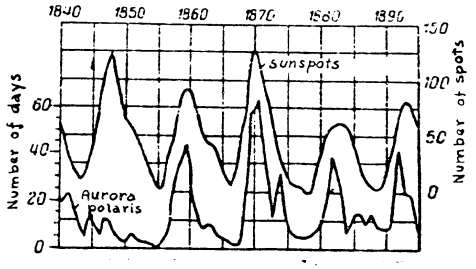

Fig.102. Number of Days in the Year with Sunspots and Aurora, Observed in Australia from 1840 to 1896.

Like magnetic activity, the aurora have a tendency to a 27-day cycle, and, fi-

nally, also have an 11-year cycle of maximum frequency. Fig.102 gives the curves of a number of sunspots (upper curve) and a number of days in the year with auroral displays (lower curve), observed during the period from 1840 to 1896.

Except for the first cycle, the maxima and minima of the curves are almost in the very same years, which indicates the close connection between these phenomena.

All this forces us to assume that the aurora, the magnetic disturbances, and the appearance of spots on the sun are interrelated.

CHAPTER VII

THEORY OF MAGNETIC VARIATIONS AND AURORA.

Section 1. The Ionosphere and its Properties.

The results of mathematical analysis of the curves of magnetograms give reason for holding that the cause of both solar and lunar diurnal variations is the existence in the upper layers in the atmosphere of a system of electric currents with a distribution that must be roughly about what is shown in Fig.73 and 74. For this reason all theories of diurnal variations, starting out from this proposition, attempts to give the mechanism of origin of these currents, under the assumption that the upper layers of the atmosphere possess a corresponding conductivity. Until the 1920's, the question of the conductivity of the upper layers of the atmosphere remained a pure hypothesis, and consequently the theories of the magnetic variations based on this hypothesis could not claim to be reliable. But observations on the propagation on the radio waves, especially short waves, showed the existence at a height of 100 to 300 km of conducting layers having the ability to reflect radio waves, as occurs with metallic conductors. In this way the hypothesis of the conductivity of the upper layers now has experimental confirmation. The existence of conductivity is explained by the ionization by the atmosphere under the action of the ultraviolet and corpuscular radiation of the sun. This region of the atmosphere has received the name of the ionosphere. Since the ionosphere plays an immense role in

the formation of the variations of the magnetic field of the earth, it is necessary, even though only briefly for us to dwell on its properties, and the methods of studying it.

1. The propagation of radio waves in the ionosphere. The dependence of the index of refraction on the density of ionization. Let us assume that in an ionized medium located in the magnetic field of the earth, there is propagated a monochromatic plane electromagnetic wave with angular velocity w in the direction of the z axis, which makes the angle B with the direction of the magnetic field H. In this case, under the action of the electric vector E of this wave, the free electrons are put into motion, and the equation of one of them will obviously be:

$$m\frac{d^2\mathbf{r}}{dt^2} = -\mathbf{E} - \frac{e}{c}\left[\frac{d\mathbf{r}}{dt}\mathbf{H}\right], \qquad (7.1)$$

if we consider that the electrons do not lose their energy under the action of the collisions, where r represents the vector of displacement under the action of E, and c is the speed of light.

If δ denotes the dielectric constant of the ionized medium, then

$$\varepsilon\mathbf{E} = \mathbf{E} + 4\pi\mathbf{P},$$

where P is the vector of electric polarization, which may in turn be represented in the following form:

$$\mathbf{P} = -Ne\mathbf{r},$$

where N is the number of free electrons in 1 cc.

From this, expressing E and r in terms of P, and substituting in equation 7.1, we obtain:

$$-\frac{m}{Ne}\frac{d^2\mathbf{P}}{dt^2} = -\frac{4\pi e}{\varepsilon - 1}\mathbf{P} + \frac{1}{Nc}\left[\frac{d\mathbf{P}}{dt}\mathbf{H}\right]. \qquad (7.2)$$

Since the vector E is a harmonic function then $\mathbf{E} = \mathbf{E}_0 \sin wt$, equation 7.2 allows us to find the value of ε connected with the index of refraction n, which is as follows:

$$\varepsilon = n^2 = 1 - \frac{2a(1-a)}{2(1-a) - b^2H_x^2 \pm \sqrt{b^4H_x^4 + 4b^2H_z^2(1-a)^2}}, \qquad (7.3)$$

where

$$a = \frac{4\pi Ne^2}{m\omega^2} \quad \text{и} \quad b = \frac{e}{mc\omega}.$$

Let us now consider the case where the magnetic field is absent, or where it may be disregarded. Then, putting H = 0 in equation 7.3, we shall have:

$$\varepsilon = n^2 = 1 - a = 1 - \frac{4\pi Ne^2}{m\omega^2}. \qquad (7.4)$$

This expression has a physical meaning only for a < 1. At a = 1, the index of refraction becomes zero, which corresponds to the phenomenon of total internal reflection. Consequently, at constant ionization density, i.e. when N = const., it is always possible to select a frequency of oscillation of the electromagnetic wave w, at which the index of refraction shall be zero. In this case the total internal reflection of the incident wave takes place, and with normal incidence it returns to the place from which it came. It is from the difference between the time of the emission of the signal and time of its arrival that the height of the reflecting layer can be determined. The frequency w_c, at which total internal reflection takes place is termed critical; if we know it we can determine the relation of the density of ionization to the mass of an ionized particle from the equation:

$$\omega_c^2 = \frac{4\pi Ne^2}{m}, \qquad (7.5)$$

whence

$$\frac{N}{m} = \frac{\omega_c^2}{4\pi e^2} = \frac{\pi f_c^2}{e^2}, \qquad (7.6)$$

where f_c is the critical frequency of the oscillations expressed in cycles per sec-

ond.

In this way the method of reflection of radio waves of variable frequency from the ionized layer allows us simultaneously to find the height of the reflecting layer and the density of ionization of that layer, provided we know the species of ionized particles. Thus if the particle is an electron whose mass $m = 9 \times 10^{-28}$ g, then

$$N = 1{,}24 \cdot 10^{-8} f_c^2.$$

If, however, the ionized particles are oxygen and nitrogen ions, the mean mass of which is $m = 3.7 \times 10^{-23}$ grams, then

$$N = 2{,}5 \cdot 10^{-4} f_c^2.$$

If an electromagnetic wave of constant frequency penetrates into an ionized layer of gas of variable ionization density, then it may reach a level at which n becomes zero. At this height, then, there is total reflection from the layer.

For the case when the magnetic field is not equal to zero, the index of refraction, as show by equation 7.3, has two values. For this reason, when an electromagnetic wave passes through an ionized medium, it must be split into two. One of these waves, corresponding to the plus sign in equation 7.3, is called the ordinary wave, while the second, corresponding to the minus sign, is called the extraordinary wave, by analogy with the doubly refracting power of crystals.

If expression 7.3 is equated to zero, then we obtain two equations each of which determines the critical frequency of the ordinary and extraordinary waves.

One of the roots of this equation $a = 1$, while the second $a = 1 \pm bH$. The first of them corresponds to the ordinary wave, the second to the extraordinary. If we replace a and b by their values and denote the critical frequency of the ordinary wave by w_c and of the extraordinary wave by w_c', then we obtain

$$\omega_c^2 = \frac{4\pi N e^2}{m}, \qquad \omega_c'^2 = \frac{eH}{mc}\omega_c' = \frac{4\pi N e^2}{m} = \omega_c^2. \tag{7.7}$$

The quantity eH/mc represents the angular velocity of rotation of the charge e about the lines of force of the magnetic field. If we denote it by w_H, we shall have

$$\omega_c'^2 \mp \omega_H \omega_c' - \omega_c^2 = 0,$$

or, passing over to frequencies expressed in cycles per second:

$$f_c'^2 \mp f_H f_c' - f_c^2 = 0,$$

whence

$$f_c = f_c' \sqrt{1 \mp \frac{f_H}{f_c'}}.$$

As shown by experience, f_H is always less than f_c, and therefore we may write, approximately:

$$f_c' - f_c = \mp \frac{1}{2} f_H = \mp \frac{1}{4\pi} \frac{eH}{mc},$$

but only the plus sign has a physical meaning, i.e. the frequency of the extraordinary wave is always somewhat higher than the freqency of the ordinary wave.

For a layer of the atmosphere the ionization of which is represented by electrons:

$$f_c' - f_c = 1{,}4\ H\ \textit{мгц}.$$

For middle latitudes, where H = 0.5 Θ_e, the difference in the frequency of the extraordinary and ordinary wave will be of the order of 0.7 megacycle.

But if the ionization is represented by ions, whose mass is a few thousand times as great as the mass of an electron, then f_H, and consequently also the difference $f'_c - f_c$, will likewise be a few thousand times as great, i.e. of the order

of a few hundred cycles.

It follows from this that in the presence of double refraction, when the difference between the frequencies of the ordinary and extraordinary rays is of the order of 1 - 1.5 megacycle, we have an ionized layer represented by electrons. In the absence of double refraction, however, the opposite conclusion cannot be drawn, since the extraordinary ray, after being reflected, may also fail as a result of absorption to reach the earth's surface.

Equation 7.7 shows that the ordinary wave coincides with a wave that is with the wave propagated in the absence of the magnetic field.

2. Methods of studying the ionosphere. The theory of propagation of radio waves in the ionosphere allows methods to be established by the aid of which it is possible to determine the height of the ionized layers, the density of their ionization and the character of the ionized particles. The principal method in the study of the ionosphere today is the pulse method.

The pulse method is as follows. A radio installation, consisting of a transmitter, receiver, antenna, pulse modulator, relaxation oscillator, and cathode oscil-

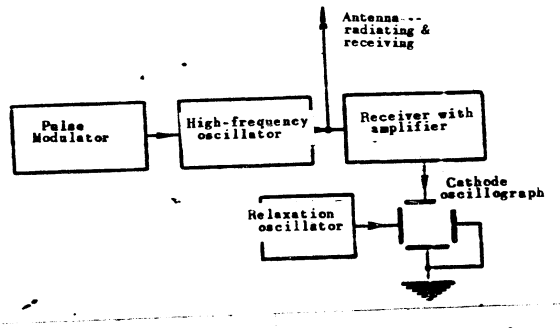

Fig.103. Scheme of the Pulse Method of Studying the Ionosphere.

lograph (Fig.103) periodically emits short pulses with a duration of the order of

10^{-4} to 10^{-5} sec. Since the frequency of the oscillations of the transmitter is of the order of a few megacycles, a few hundred oscillations will go into each pulse. The time interval between two successive pulses, or in other words the pulse frequency, is so chosen that the radiated impulse shall be able to reach the reflecting layer and return from it before the second impulse begins. With a height of the F_2 layer of the order of 300 km, the travel time t will be 2×10^{-3} sec. Usually the interval between the pulses is taken ten times as great, i.e. 1/50 sec. The direct and reflected pulses are received by the antenna and enter the radio receiver, at the output of which an alternating voltage is produced. This voltage is fed to one of the pairs of plates of a cathode oscillograph, while the other pair is connected to the relaxation oscillator of the same frequency as the pulse modulator, 50 cycles.

The relaxation oscillations are synchronized with the pulse modulator in such a way that at the time of minimum amplitude of the relaxation oscillations, i.e. when the voltage across the oscillograph plates is at a minimum, the oscillators shall receive a pulse. In that case the reflected pulse arriving after a certain time interval and entering the oscillograph plates, meets the cathode beam deflected in a direction perpendicular to the direction of the oscillations of the pulse, and on the oscillograph screen we shall see two luminous bands a certain distance apart. Since the relaxation oscillations are proportional to time over most of its half-period, the deflection of the reflected beam will likewise be proportional to its time lag. In this way, by measuring the distance between the luminous bands, if we know the time scale, the height of the reflecting layer can be determined.

To photograph these bands, the oscillograph screen is covered by an opaque screen with a narrow slit in the middle, and the photographic film is displaced perpendicularly to this slit. If the height of the reflecting layer is constant, then a series of straight lines, correspondent to a different number of reflections, will be obtained on the film. But if the height varies, we shall obtain a curve showing the relation of the height of the layer to the time.

STAT

The pulse method allows us to determine only the height of the reflecting layer corresponding to a given frequency of the transmitter. To find the critical frequencies by this method, the pulse observations and measurements are made with a smooth variation of the transmitter and receiver frequencies from low frequencies up to the frequency which the reflection stops. In modern installations, these variations of frequency are produced automatically, and such an installation has received the name of ionosphere station.

3. Measurement of the height of the reflecting layer. If the velocity of propagation of waves in the ionosphere did not differ from the velocity of propagation in a neutral medium, i.e. if it were equal to the speed of light c, then the height h_g of the reflecting layer could be calculated by the formula

$$h_g = \frac{1}{2} c\tau,$$

where τ is the time interval between the moment of emission and the moment of arrival of the reflected wave. In view of the existence of dispersion, the velocity for each frequency in the ionosphere will be different and will vary with the variation of the ionization density. For this reason the true height h will differ from the height h_g termed the effective height, and must be calculated by the formula:

$$h = h_0 + \int_{h_0}^{h} u\,dt = h_0 + \int_{h}^{h} dh, \qquad (7.8)$$

where h_0 is the height of the beginning of the layer above the earth's surface and is the velocity of propagation of the waves in the ionosphere layer.

By means of appropriate transformations, equation 7.8 is brought into the following form:

$$h = \frac{2}{\pi} \int_0^{\omega} \frac{h_g\, d\omega_1}{\sqrt{\omega^2 - \omega_1^2}}.$$

Since w_1 is an arbitrary variable, it may be taken as equal to w sin φ, i.e.

$$\omega_1 = \omega \sin \varphi, \qquad (7.9)$$

and in that case

$$h = \frac{2}{\pi} \int_0^{\pi/2} h_g \, d\varphi,$$

where h_g is a function of w sin φ. Consequently the true height is numerically equal to the area between the curve $h_g = f(\varphi)$ and the axis of abscissas. Since the photograms give us h_g as a function of w, it follows that to find h_g as a function of φ, must be calculated by equation 7.9 for the given frequency w and for various values of w_1, and the values of h_g corresponding to the values of φ must be found from the photogram, and this will in fact give us h_g as a function of φ.

If we now construct a graph of h_g against φ and integrated by a planimeter, we shall find h for the given frequency w.

Observations of the reflections of the radio waves by the critical frequency method, that is, the determination of the frequency at which the reflected wave vanishes, have allowed the existence of three ionized layers at mean heights of 100, 200 and 300 km to be established. These are termed respectively the E, F_1 and F_2 layers.

The critical frequencies corresponding to these layers depend both on the latitude and longitude of the place and on the solar activity and as well as the solar altitude. Table 22 gives the values of the mean annual critical frequencies at noon when the critical frequencies are at maximum, as well as the number of sunspots in the respective years.

The frequencies are expressed in megacycles and are given for various points of the earth, as well as for various epochs within the limits of a single cycle of solar activity.

As indicated by the Table, the critical frequency, and consequently also the

ionization density in all layers increases with the solar activity and with decreasing latitude of the places.

Table 22.

Mean Annual Critical Frequencies and Number of Sunspots.

Point	Latitude φ	Longitude λ	Year	Number of sunspots	Critical frequency at noon		
					E	F_1	F_2
Washington . . .	38°50′ N	77°0′ W	1933	6	3.03	3.34	—
			1934	9	3.10	3.28	5.71
			1935	36	3.35	3.49	6.45
			1936	80	3.72	4.34	9.34
			1937	114	3.75	4.61	10.0
			1938	110	3.74	5.21	10.05
			1939	84	3.73	5.26	9.57
			1940	69	3.52	4.58	8.42
			1941	46	3.37	4.43	7.29
Huancayo, Peru .	22°3′ S	75°20′ W	1939	84	4.04	5.46	10.36
Watheroo, Australia .	30°19′ S	115°53′ E	1939	84	4.44	5.16	10.52
Tomsk	56°30′ N	84°54′ E	1939	84	3.3	5.2	9.5
Tromso, Norway	69°40′ N	18°95′ E	1939	84	3.07	4.87	7.33

In exactly the same way, the ionization of the F_1 and E layers, as shown above, depends on the zenith distance of the sun, and therefore has a diurnal and seasonal march. Thus, in the course of a day, the ionization increases from midnight, reaching a maximum at noon, and then symmetrically declines to reach a minimum at midnight again.

The F_1 and E layers are observed only in the daylight hours and at latitudes where the zenith distances of the sun is less than 70°. At great zenith distances, i.e. in the evening and night hours, and also at high latitudes, the F_1 layer disappears, and merges with the F_2 layer.

The region in which layers are observed is termed the F region.

By using equation 7.6 and the data of Table 22, is easy to obtain the value of the ionization density in the E, F_1 and F_2 layers. Thus, if the ionized particle is

considered to have the mass of an electron, then its mean ionization is of the order $N = 1.5 \times 10^5$ for the layer; $N = 3.0 \times 10^5$ for the F_1 layer; and $N = 10 \times 10^5$ for the F_2 layer.

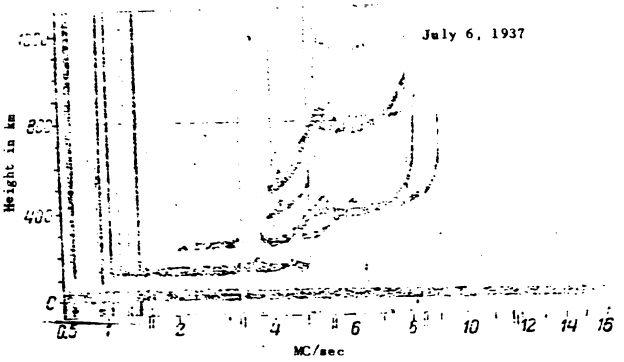

Fig.104. Photogram of Height-Frequency Characteristic with Reflections from E and F Layers.

The fact of double refraction in the F_1 and F_2 layers and its absence, in most cases, in the E layer is of importance, for it gives us grounds to hold that the F_1 and F_2 layers consists of electrons while the question of the composition of the ionized particle in the E layer still remains open, since double refraction in the layer is met very seldom, excepting in the high layers.

Fig.104 shows a typical photogram of the height-frequency characteristic with a reflection from the E layer, and from the F_1 and F_2 layers, obtained by means of an automatic record of the reflected waves of various frequencies. The frequencies are plotted along the abscissa axis in megacycles, and the effective heights in kilometers on the ordinate axis. As will be seen, the low frequencies are reflected from the lower layer at a height of the order of 100km (the E layer); at a frequency of about 400 KZ the reflection is interrupted and begins again from a higher lay-

er at a height of about 150-160 km (the F_1 layer). With increasing frequency, the height of the layer increases, and at a frequency of about 5000 KZ a new reflection corresponding to the F_2 layer at a height of 280 km, begins. At a frequency of over 6000 kz, the height of the layer increases sharply, and the curve, by bifurcating, moves upward into branches, and the reflection stops.

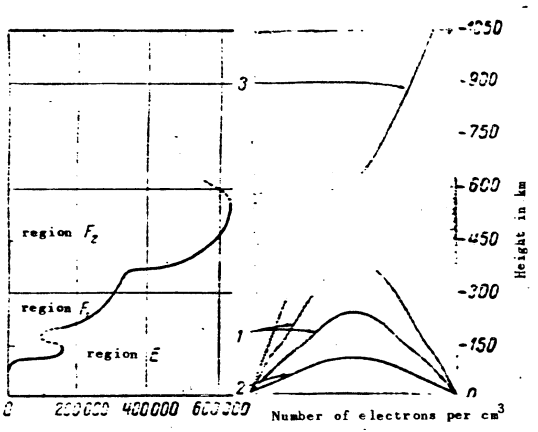

Fig.105. Density of Ionization at Various Heights, and the Paths of Rays of Various Frequencies.

1. path of waves of medium frequency; 2. path of low frequency waves; 3. path of high frequency waves.

The bifurcation of the curve corresponds to the reflection of the ordinary and extraordinary wave. The upper, fainter curve is the result of being twice reflected, as a result of which its ordinate, i.e. its height, is double that of the lower curve.

Fig.105 gives a graph of the ionization density against height obtained by working up a photogram, and indicates the paths of rays of various frequencies emit-

ted from one point of the earth's surface and received at another point.

4. Composition of the ionosphere and formation of ionized layers. On the basis of data on the spectra of the aurora and luminescence of the night sky, it will be assumed that at heights 100 - 120 km (the E layer) that the ionosphere consists of molecular nitrogen N_2 and molecular oxygen, O_2. Above 120 km, the dissociation of the oxygen begins, and in atomic oxygen, O_1, begins to predominate in the composition of the ionosphere. For this reason it is usually considered that the E layer consists of ionized nitrogen, N_2, when it bifurcates into two layers, layer F_1 consists of N_2 while layer F_2 consists of O_1. In higher regions dissociation of nitrogen begins and atomic nitrogen, N_1 begins to play a role.

An entirely definite energy (the ionization potential) is required for the ionization of oxygen and nitrogen molecule or atoms. This energy has the following values expressed in electron volts ($1eV = 1.59 \times 10^{-12}$ erg):

gas	O_2	O_1	N_2	N_1
energy	12.5	13.5	15.8	14.5

This energy may be obtained from solar radiations which consist of two forms: corpuscular and waves. The role of the corpuscular radiation, however is very insignificant and is manifested only at moments of enhanced solar activity. On normal days, however, as shown by research, the principal ionizing factor is the wave radiation in its shortwave portion. Indeed, from the condition that the energy of a light quantum $h\nu$ must be not less than the value of the ionization potential above mentioned, it follows that ionizing radiation must have a wavelength between 850 and 1000 A.

Alongside of ionization the reverse phenomenon also takes place, namely the recombination of ions, which proceeds faster the greater the density of the atmosphere. At a height of the order of 100 km, the equilibrium state between the processes of ionization and recombination is rather rapidly established, and therefore the density of ionization (the number of free electrons per cc) at these heights must be a function of the solar altitude. At great heights, owing to the low density of the at-

mosphere, there is a lag in the processes of recombination, as a result of which the ionization is shifted in phase with respect to the solar altitude.

Since the density of the atmosphere decreases with increasing height, while the ionizing radiation cannot penetrate into the low layers, the ionization density must reach a maximum at a certain height, leading to the layered structure of the ionosphere. The existence are not of a single layer but of several, is explained by the complex composition of the atmosphere and by the different height distribution of different elements.

The principal ionization-recombination equation determining the state of the ionosphere has, as is well known, the following form:

$$\frac{dn_e}{dt} = q - q', \qquad (7.10)$$

where n_e is the density of ionization, q the number of electrons newly formed per second under the action of the ultraviolet radiation, and q' the number of electrons recombining per second.

Obviously the value of q is proportional to the intensity of the incident monochromatic radion W, that is, to the quantity of energy passing per second through 1 cm^2 of surface and normal to the radiation, and to the number of neutral particles $n - n_e$, and is inversely proportional to the quantity of energy w_0 which is necessary for a single event of ionization, i.e.

$$q = \tau(n - n_e)\frac{W}{w_0}, \qquad (7.11)$$

where τ is the coefficient of photoabsorption, n the number of neutral particles per cc in the un-ionized atmosphere, and n_e the number of electrons per cc.

The number of q is proportional both to the number of electrons n_e and the number of positive ions n_+, i.e.

$$q' = \varkappa n_e n_+,$$

where $\varkappa$ is the coefficient of electronic recombination.

Since $n_e \approx n_+$, then

$$q' = \alpha n_e^2.$$

equation 7.10 may thus be written in the form

$$\frac{dn_e}{dt} = q - \alpha n_e^2. \qquad (7.12)$$

In addition to the direct processes of photoionization of electron and of their recombination, a considerable number of other processes must also exist, as a result of which electrons appear and disappear; they in these processes include: the adhesion of electrons to a neutral particle, leading to the formation of negative ion; the separation of electrons from negative ions; and the recombination of ions. In reality, therefore, equation 7.12 is more complex in form, but, as shown by calculations, it may be brought into the same form 7.12 if alpha is taken to mean a certain effective coefficient of recombination depending on the coefficient of adhesion, separation, and recombination of ion.

Ya. L. Al'pert gives the following values of q, n_e and alpha for various layers of the ionosphere:

Layer	E in summer	F_1 in summer	F_2 in winter
n_e (electrons) per cc	1.5×10^5	3×10^5	2.5×10^6
a (cc/sec)	10^{-8}	3×10^{-9}	1.5×10^{-10}
q	200	300	800

These values relate to a time close to noon, to maximum solar activity, and measurements in the middle latitude of the northern hemisphere.

To find the height of the layer where ionization is maximum, let us now return to equation 7.11 and substitute in it, for the quantity W, the intensity of radiation beyond the boundaries of the ionosphere. For this purpose, let us consider the passage of radiation at the angle z to the vertical through a plane layer of at atmosphere of thickness dh. The losses of energy in this case will be

$$dW = z(n - n_e) W \sec z \, dh.$$

On integrating this equation between the limits from height h to infinity, we obtain:

$$W = W_\infty e^{-z \sec z \left[\int_h^\infty n \, dh - \int_h^\infty n_e \, dh\right]}.$$

If we substituted this value in the equation 7.11 we have

$$q = z \frac{W_\infty}{\omega_0} (n - n_e) e^{-z \sec z \left[\int_h^\infty n \, dh - \int_h^\infty n_e \, dh\right]}. \qquad (7.13)$$

The integrals in the exponent of the power of e represent the total number of neutral particles and electrons in a column 1 sq cm^2 cross section extending in height from level h to infinity.

The maximum number q_m and the height h_m at which this formation takes place the finding of the maximum number q_m and the height h_m at which such formation takes place is very much simplified if it is assumed that the layer is almost un-ionized, i.e. if we neglect the value of n_e.

In that case:

$$q = z \frac{W_\infty}{\omega_0} n e^{-z \sec z \int_h^\infty n \, dh}. \qquad (7.14)$$

The number of neutral particles, n, in this equation, is determined by the barometric formula

$$n = n_0 e^{-ah}, \qquad (7.15)$$

in which the constant **a** has the value:

$$a = \frac{\mu g}{RT},$$

where μ is the molecular weight of the gas, **R** the gas constant, **g** the acceleration of gravity, and **T** the absolute temperature.

The height h_m at which **q** reaches its maximum value, is found from the equation

$$\frac{dq}{dh} = 0.$$

This equation gives

$$h_m = \frac{1}{a} \ln \frac{\tau n_\infty \sec z}{a}. \qquad (7.16)$$

It follows from equation 7.15 that at this height **n** takes the value:

$$n_m = \frac{a}{\tau \sec z}.$$

on substituting the expressions for h_m and n_m in equation 7.14, we obtain the value of the maximum of **q**:

$$q_m = \frac{a W_\infty}{\alpha_1 e \sec z}.$$

Equation 7.16 shows that the maximum formation of photoelectrons at a given zenith distance **z** is determined by the value of τ and the molecular weight μ of the ionized gas, since the coefficient **a** depends at constant temperature only on μ.

To find the height h_m of the ionized layer at which the maximum formation of electron q_m takes place, the term containing n_e in equation 7.13 must also be taken into account. In that case the condition of the maximum $\frac{\partial q}{\partial h} = 0$ yield:

$$\frac{\partial n}{\partial h} - \frac{\partial n_e}{\partial h} = -\tau (n - n_e)^2 \sec z,$$

whence as it is from this that we find the value of h_m.

But the maximum number of electrons formed is still insufficient for the appear-

ance of a layer with maximum electron density n_e. For this it is necessary that the disappearance, i.e. the process of recombination, shall not compensate the maximum n_e. The existence of layers with a maximum of density shows that such compensation in reality does not exist.

To find the height of the layer the maximum density of ionization it is necessary to solve the equation 0 with respect to h, but for this we must know the function of n_e, and this is found from the ionization-recombination formula of equation 7.12.

It must be borne in mind that the maximum n_e must not by any means coincide with the maximum of q.

But the solution of equation 7.12 meets with difficulties, since up to now the numerical value of the coefficients characterizing one process or another of ionization and recombination has not yet been established. As exactly the same value of the coefficient τ has been established on the basis of quantum calculations, only for atomic oxygen O_1.

5. The conductivity of the ionosphere.

Thanks to the presence of free N ions in the atmosphere, it becomes electrically conductive, and by the electronic theory, its conductivity, γ, is expressed by the relation:

$$\gamma = Ne^2 \frac{l}{2mv}, \qquad (7.17)$$

where l is the mean free path of a particle, v the mean velocity of the thermal motion of the particle, and m its mass. But if the ionized layer is in a magnetic field, then the conductivity in the direction of the field H does not remain the same as in the absence of the field, and in the direction perpendicular to the field, the conductivity γ' now becomes equal to

$$\gamma' = \frac{\gamma}{1 + \left(\frac{l}{r}\right)^2}. \qquad (7.18)$$

where r is the radius of eddying vertical motion of the electrons about the lines of force of the magnetic fields H, determine the equation

$$r = \frac{mv}{He}. \qquad (7.19)$$

The mean free path of a molecule is determined, according to the kinetic theory of gases, by the relation

$$l\sqrt{2}\pi nd^2 = 1, \qquad (7.20)$$

where n is the number of molecules for 1 cc cm^3 and d the diameter of those molecules.

Table 23 gives the values of n and l for various heights, calculated from equation 7.20 and the barometric formula 7.15.

Table 23

Number of Molecules and Mean Free Path at Various Heights

h, km	60	100	140	200	300
n	9.9×10^{15}	1.6×10^{4}	4.4×10^{12}	2.5×10^{8}	7.7×10^{6}
l, cm	0.03	1.5	57	10^{1}	10^{7}

The radius of vertical motion of the electron depends on the magnitude on the magnetic field strength, and therefore, for particles of the moving at the same velocity, they will differ at different latitudes. It takes the greatest value at the magnetic equator by H = 0.3 O_e. Therefore, for electrons on the equator, r = 1.3 cm, while for ions r = 20 cm.

Consequently the ratio l/r for electrons takes greater values for height over 100 km, and for ions, for heights over 150 km.

Table 24 gives the values of the specific conductivities of various layers, calculated by formulas 7.17 and 7.18 in the CGSμ system.

It was assumed for this calculation that the velocity of thermal motion of particles v corresponds to the absolute temperature T = 360°, which, as shown by ra-

dio measurements, the E layer does have.

Table 24

Values of the Specific Conductivity in Various Layers

LAYER	N/m	$\gamma_{e\,a}$	$\gamma_{m\,a}$	$\gamma_{e\,\perp}$	$\gamma_{m\,\perp}$
E . .	$0.87 \cdot 10^{12}$	$3,66 \cdot 10^{-14}$	$1.25 \cdot 10^{-11}$	$0.33 \cdot 10^{-14}$	$0,62 \cdot 10^{-11}$
F_1 . .	$1.6 \cdot 10^{12}$	$3.3 \cdot 10^{-14}$	$2,48 \cdot 10^{-11}$	$\sim 10^{-23}$	$\sim 10^{-23}$
F_2 . .	$6.4 \cdot 10^{12}$	$12,2 \cdot 10^{-14}$	$9.1 \cdot 10^{-11}$	$\sim 10^{-23}$	$\sim 10^{-23}$

At greater heights the temperature in all probability is higher, but the experimental data on the temperature of the upper layers is still inadequate, and therefore it has been taken at one and the same for all layers. This temperature corresponds to a velocity $v = 6.8 \times 10^6$ cm/sec.

6. The sporadic layer.

Besides the E, F_1 and F_2 layers a layer is often observed at the height of the E layer with a critical frequency exceeding the critical frequency of the E layer, and sometimes even the critical frequency of the F_2 layer, which points to the existence of reasons with an enhanced degree of ionization in the E layer. Since this layer is not found everywhere, nor is found at different times of the day and year, it has been termed the sporadic layer or E_s layer. The sporadic layer has been found to appear most often in the middle and high latitudes and to have a local character. In the middle latitudes the E_s layer appears more often in summer, from May to September.

Observations show that the waves reflected from the sporadic layer contains both an ordinary and an extraordinary wave. This gives grounds for assuming that the sporadic layer has an electronic structure, but the nature of its origin has not yet been elucidated.

7. Tidal phenomena in the ionosphere. Observations on the reflection of radio signals from the E layer have allowed the establishment in this layer of the existence of semidiurnal fluctuations at height corresponding to the tidal phenomena caused by the moon. The amplitude of such fluctuations, as shown by the observations, reach 1 km, and the maximum occurs 3/4 hour before the superior combination of the moon, as will be seen from Fig.106, which represents the results of direct

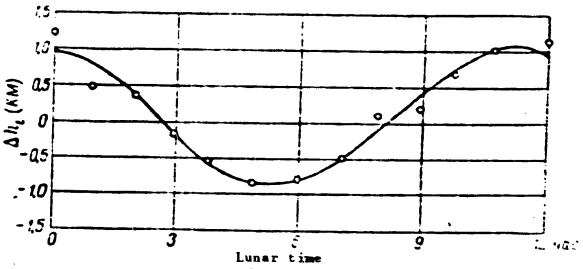

Fig.106. Semidiurnal Fluctuations of the Equivalent Height of the E Layer.

observations of the fluctuations of the E layer.

On the earth's surface tidal phenomena of the atmosphere caused by the moon are also observed, but these phenomena are manifested in the form of semidiurnal fluctuations of the atmospheric pressure p with amplitude of about 0.0000115p.

If the fluctuations of height in the E layer with amplitude Δh = 1 km are converted into fluctuations of pressure, then, according to the barometric formula of equation 7.15, we obtain

$$\Delta p = -pa\,\Delta h = 0{,}08\ p,$$

i.e. 7000 times greater than the relative fluctuations at the earth's surface.

Recent investigations show the existence of such fluctuations not only of height but also of the critical frequency f, i.e. the density of ionization. Fig. 107 shows the semidiurnal tidal fluctuations of the equivalent height and critical frequency in the F_2 layer. It will be seen that the amplitude of the fluctuations

of height in the F_2 layer are twice those in the E layer, and that in general, as shown by observations, the amplitude of both Δh, and Δf increase with the height of the layer.

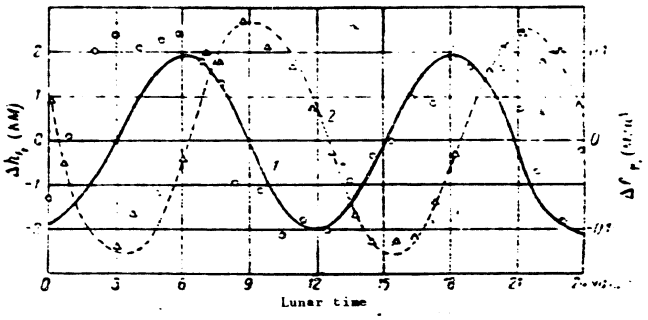

Fig.107. Semidiurnal Fluctuations of Equivalent Height in the E Layer and of Critical Frequency in the F_2 Layer.

1. fluctuations of the height of the layer; 2. fluctuations of frequency.

The lunar tidal motions of the atmosphere may be regarded as the propagations of plane waves in a medium with a certain coefficient of refraction. If such waves at a certain height, meet a region of the atmosphere where the index of refraction is equal to zero or has become negative, then these waves will be reflected, and an oscillatory motion will be set up between the earth's surface and this region. The region of the temperature maximum, at a height of the order of 15 - 30 km is believed to be such a reflective region for oscillations at the earth's surface.

To explain the lunar fluctuations in the E layer, it is assumed that in addition to the second region of temperature minimum at height 80 km, there is also a third region at height 140 - 160 km, and the fluctuations here take place between the second and third regions.

The existence of fluctuating motions in the ionosphere is very important for

the construction of the theory of the diurnal variations of the magnetic field of the earth, since one of the possibilities of explaining such variations is the motion of ionized layers in the earth's magnetic field.

Section 2. Theories of the Diurnal Variations.

The existence of the conducting layers of the atmosphere at heights of 100 to 300 km gives ground for asserting confidently that cause of the diurnal variations, and, in general, of all the magnetic variations, are the electric currents in these layers. But to explain the causes of these currents we still have no well established fact which would allow us to take one explanation or another and consider it correct. All the existing theories of the diurnal variations thus reduce to a choice of the mechanism of the origin of the current, each of which may claim a more or less degree of authenticity.

Up to now there have been three principal theories of the diurnal variations: the theory of the atmospheric dynamo proposed in 1872 and worked out in detail somewhat later (1889), the diamagnetic theory (1928), and the theory of drift currents (1929).

Recently, in connection with the work of I. Ye. Tamm (cf.infra) the diamagnetic theory has already lost its importance, while the theory of the atmospheric dynamo, as a result of the periodic motions movement discovered in the ionosphere, has acquired a higher degree of credibility.

The theory of drift currents, thanks to the same work of Tamm, has found justification and may, on the same level as the theory of the atmospheric dynamo claim a certain standing of credibility. It is therefore necessary to dwell on each of these theories and to point out their shortcomings.

1. The theory of the atmospheric dynamo. This theory, proposed by Stewart Shuster (Bibl.55) is based on the principle of the induction of electomotive force on the motion of a conductive in the earth's magnetic field in a manner similar to what takes place in a dynamo, and for this reason the theory is called that of the

atmospheric dynamo. Since the high layers of the atmosphere possess conductivity, it follows that when they move in the magnetic field of the earth H, an electric field perpendicular to H and to its direction of motion is produced, and its strength is determined by the Faraday law:

$$\mathbf{E} = [\mathbf{uH}],$$

where u is the velocity of motion of the ionized layer. Under the influence of the field E a current whose direction is perpendicular to u and H horizon in the atmosphere, while the density j is determined by the equation

$$\mathbf{j} = \gamma' \mathbf{E},$$

where γ' is the conductivity determined by equation 7.18, since j is directed perpendicularly to H.

By substituting for γ' its value taken from equation 7.18, we obtain:

$$\mathbf{j} = \frac{Ne^2}{2m\nu} \cdot \frac{l}{1 + \left(\frac{l}{r}\right)^2} [\mathbf{uH}].$$

Shuster assumed that the velocity u was due to the horizontal displacement of the upper layers of the atmosphere under the influence of the diurnal fluctuations of atmospheric pressure observed on the surface of the earth.

Barometric observations show that the pressure has two fluctuations during the course of the day: a regular one with a semidiurnal period and an amplitude of 1 mm Hg and a less regular one with a diurnal period and an amplitude of 0.3 mm. The causes of these fluctuations may be either the tidal or the thermal effect. The air currents leaving the high pressure regions located in the meridian at the equator, form electric currents of the same character, as shown in Fig.73. During the daytime the currents moving northward in the vertical magnetic field Z, create currents directed westward, which as a result of the reduction in conductivity in the atmosphere with increasing latitude lead to the formation of closed currents causing the diurnal variations of the magnetic field.

Investigations have shown that the electric currents obtained as a result of the motion of the atmosphere owing to the diurnal variations of barometric pressure, although they do in their character coincide with those calculated on the basis of magnetic observations, are still markedly different from them in phase. Thus the extreme values of the northern component are observed from 10 to 11 hours (Fig.69) while according to this theory they should occur between 14 and 16 hours.

Such a disagreement may be explained by the horizontal displacements in the upper layers due to the fluctuations of the barometric pressure are assumed to be the same as the displacements observed on the surface of the earth, which might not in reality be the case at all.

The contradiction with theory was found in the value of the conductivity. If the velocity u is considered a consequence of the diurnal barometric fluctuations, the amplitude of which amounts to 1 mm, then for the equator it should have a value of the order of 30 cm/sec.

On the basis of the current distribution map (Fig.73) the strength of the current flowing through a cross section of the layer 1 cm wide and with the width equal to the thickness of the layer is equal to 3×10^{-5} CGSμ, while the vertical component of the magnetic field Z at latitude 30° equal to 0.3 O_e, and thus the conductivity of a vertical column of the ionosphere 1 cm^2 in cross section and with the height equal to the thickness of the layer is determined by the equation:

$$\gamma = \frac{3 \cdot 10^{-5}}{30 \cdot 0.3} = 3 \cdot 10^{-3}\ CGS\mu.$$

The thickness of the E layer is of the order of 50 km and therefore the specific conductivity of the E layer should be

$$\gamma \sim 10^{-11}\ CGS\mu.$$

On comparing this value with the values for the conductivity in Table 24 we see that they agree with the value of the ionic conductivity, but are a thousand times smaller than the electronic conductivity. For this reason, in order to reconcile the

theory with the observations it would be necessary to assume either the existence of a predominance of ions in the E layer, or the existence of velocities of displacement of the ionized masses of the atmosphere that are many times as great as the velocities observed at the surface of the earth (30 cm/sec).

At the present time neither of these possibilities are excluded, since occasionally we do succeed in observing the splitting of a reflected ray in the E layer into two, an ordinary and an extraordinary, with a frequency difference corresponding to the ionic nature of the layer. But the existence of fluctuations of pressure in the E layer 7000 times as great and the fluctuations at y the surface of the earth give grounds for considering that the velocity of displacement must also be many times greater than 30 cm/sec.

2. The theory of the drift current (Bibl.56). According to this theory the charged particles in the region of long free paths, moving in the magnetic and gravitational fields of the earth, must experience a translatory motion (drift) in a direction perpendicular to the force of gravity mg and to the magnetic field H. The velocity of this drift would be:

$$u = \frac{mg}{eH} \sin \varphi, \qquad (7.21)$$

where Φ is the angle between H and the force of the gravity, or the vertical. Here H and e must be expressed in absolute electromagnetic units.

Since the magnetic field of the earth may be considered in first approximation as the field of an elementary magnet, its field strength H may be represented in the form:

$$H = H_0 \sqrt{1 + 3 \sin^2 \varphi}, \qquad (7.22)$$

where H_0 is the field strength at the equator and Φ is the latitude.

Moreover, for an elementary magnet the following relation holds true:

$$\operatorname{ctg} \varphi = 2 \operatorname{tg} \psi.$$

For this reason, expressing sin Φ in terms of tan Φ in equation 7.21 and substituting tan Φ according to the following formula, we shall have:

$$u = \frac{mg}{eH} \cdot \frac{\cos \varphi}{\sqrt{1 + 3 \sin^2 \varphi}}.$$

On substituting for H its expression in equation 7.22, we obtain:

$$u = \frac{mg}{eH_0} \cdot \frac{\cos \varphi}{1 + 3 \sin^2 \varphi}.$$

Thence the velocity of drift of the electrons at the pole, $u_{90} = 0$, and at the equator $u_o = 1.2 \cdot 10^{-4}$ cm/sec, while the drift of the ions of the atmosphere, the mass of which $m = 5 \times 10^{-23}$ grams, is zero at the pole and 10.2 cm/sec at the equator.

The current density is obviously expressed as:

$$j = eNu = \frac{mgN}{H_0} \cdot \frac{\cos \varphi}{1 + 3 \sin^2 \varphi} CGSμ.$$

Since both electrons and ions participate in the formation of the drift currents, it follows that if the number of ions is considered equal to the number of electrons, and the density of ionization in the upper layers $N = 10^{-6}$, we have for the density of the drift current on the equator the value $j = 1.7 \times 10^{-13}$ CGSμ.

Such a drift, or, to put it differently, displacement of charged particles, produces a current in the direction from west to east, since the magnetic field is directed from south to north. The drift of positive ions to the east causes the accumulation of positive charges on the morning side, while the drift of negative ions causes their accumulation on the evening side. As a result, in the region of long free paths an electrical field arises which is directed toward the west and which, however, may be electrically connected with the lower region of short free paths, since the ions may move freely under the influence of the force of gravity along the lines of force of the magnetic fields, especially around the pole. For this reason, in the lower part of parts currents arise which with a corresponding distribution of activity may form a system of currents similar to the one shown in Fig.73. The cal-

culation of the conductivity necessary for the formation of drift currents corresponding to the calculated system of currents shown in Fig.73 indicates that the conductivity may be considered considerably lower than that required by the diamagnetic theory.

The theory of drift currents, however, encounters obstacles in explaining the forenoon extreme extremum of the northern and eastern components, and also in explaining the intensity of the observed variations. In addition some authors have expressed doubts that drift currents could lead to the formation of so complex a system of currents as the theory requires.

I. Ye. Tamm (Bibl.57) who has given a criticism of the drift current theory and for the first time pointed out the complete unsoundness of this theory if we start out from the same propositions that were established by Chapman. According to Chapman's belief, in an equilibrium system, ionized gas consists of an aggregate of ions which are free and independent in their motion and which may, in the absence of a magnetic field, fall downwards under the action of the force of gravity. In reality, however, in the equilibrium state of the atmosphere, the own weight of the ions is balanced by the partial pressure of the ion gas p', which is defined by the equation

$$p' = kTN,$$

where T is the absolute temperature, N the density of ions, i.e. their number in unit volume, and K is the Boltzman constant. For this reason, in the equilibrium state of the atmosphere, when the density of the ion is determined by the barometric formula, there can be no drift, and, consequently, the Chapman theory, just like the diamagnetic theory, is a simple misunderstanding.

But in this same work, Tamm also pointed out the possibility of retaining the drift current theory explains the diurnal variations if we start out not from the equilibrium condition of the ionosphere, but from the existence in it of a "nonuniform" distribution of densities, that is, of densities varying otherwise than by the

barometric formula. Such a distribution, in Tamm's opinion, corresponds more to the real conditions of the state of the ionosphere, which also makes possible the existence of drift currents analogous to those of Chapman. Tamm gives a derivation of the formula for the density of the drift current on the basis of a statistical model of the ionosphere obtained by solving the Boltzman kinetic equation. Since in its general form the Boltzman equation is very complex, Tamm introduced two assumptions to simplify it: first that the distributions of ions and electrons by velocities in each element of volume of the ionosphere differs little from the Maxwell distribution corresponding to the temperature T, and secondly that the ordered motion of ions of the atmosphere is quasi stationary, that is, at each given moment the motion is determined by the instantaneous distribution of the ion density and the temperature, as a result of which the derivatives of density and temperature with respect to time may be neglected and the density and temperature considered assigned functions of the coordinates.

Under such assumptions, the kinetic equation of the stationary motion of each species of ion will have the form:

$$\mathbf{v}\,\mathrm{grad}\,f + \left(\mathbf{g} + \frac{e}{m}\mathbf{E} + \frac{e}{mc}[\mathbf{vH}]\right)\mathrm{grad}_v f = Pf, \qquad (7.23)$$

where v is the velocity of motion of the ions, g the acceleration of gravity, E the electric field strength, H the magnetic field strength, e/m the charge-ratio, f a function of the distribution of ions, which differs from the Maxwell function by the small quantity Φf_0, i.e.

$$f = f_0(1 + \Phi),$$

while

$$f_0 = N\beta^{3}\pi^{-\frac{3}{2}}e^{-\beta^2 v^2}, \quad \text{a} \quad \beta = \frac{m}{2kT}.$$

The term P_f, taking account of the collisions between the ions themselves and with neutral and between ions and neutral molecules, is assumed to be proportional

to the deviation of the distribution function f from its equilibrium value f_0, i.e.

$$Pf = -\frac{v}{\lambda}(f - f_0) = -\frac{v}{\lambda} f_0 \Phi.$$

where λ is the length of the free path.

After substituting the values f, f_0 and P_f, equation 7.23 now takes the form:

$$\left(\frac{e}{mc}[\mathbf{vH}]\,\mathrm{grad}_{\mathbf{v}}\,\Phi\right) + \frac{v}{\lambda}\Phi = (\mathbf{vq}),$$

where

$$\mathbf{q} = \beta^2(2\mathbf{a} - \beta^2 v^2 \mathbf{b});$$

$$\mathbf{a} = \mathbf{g} + \frac{e}{m}\mathbf{E} - \frac{1}{2\beta^2}\,\mathrm{grad}\,\ln\left(\frac{N}{T^{3/2}}\right). \qquad \mathbf{b} = \frac{\mathrm{grad}\,T}{2T}.$$

The solution of this differential equation is in the following form:

$$\Phi = \frac{\lambda\beta}{s^2 + \beta^2 v^2}\left[\beta v(\mathbf{vq}_\perp) + \frac{s}{H}\mathbf{v}\,[\mathbf{qH}] + \frac{\lambda}{v}(\mathbf{vq}_\parallel)\right],$$

where $\mathbf{q}_\perp$ is the vector perpendicular to the magnetic field H, $\mathbf{q}_\parallel$ is the vector parallel to that field, and H is the ratio between the free path and the diameter of the circular orbit of a particle in the magnetic field, i.e.

$$s = \frac{\lambda}{2r} = \frac{e\lambda H}{2c\sqrt{2mkT}}.$$

The density of the current formed by the motion of charges of one sign is determined by the laws of statistics as:

$$\mathbf{j} = e\int \mathbf{v} f_0 (1 + \Phi)\, dv_x\, dv_y\, dv_z.$$

Integration of this expression on substitution of the value of Φ is possible only in the two limiting cases when $\lambda \ll 2r$ or also, when λ is $\gg 2r$. The former case represents the motion of ions in the lower ionosphere, while the second case represents that motion in the upper ionosphere.

For the case $\lambda \ll 2r$, we obtain:

$$j_{\perp} = \frac{4\pi N^2 \sqrt{\pi}}{3\sqrt{2kT}} \left(g + \frac{e}{m} E - \frac{kT}{m} \cdot \frac{\operatorname{grad} N}{N} - \frac{k}{2m} \operatorname{grad} T \right).$$

For the case, however, where $\lambda \gg 2r$:

$$j_{\perp} = \frac{c}{H^2} [(Nmg + NeE - \operatorname{grad}(kT.N)] H].$$

Since the total current density is equal to the sum of the currents of positive and negative ions, i.e.

$$j = j_{+} + j_{-},$$

then

$$\left.\begin{aligned} j_{\perp} &= \frac{4\pi N^2}{3\sqrt{2kT}} \left(\frac{\lambda_{+}}{\sqrt{m_{+}}} + \frac{\lambda_{-}}{\sqrt{m_{-}}} \right) E \quad (\lambda \ll 2r) \\ j_{\perp} &= \frac{c}{H^2} [(Nmg - \operatorname{grad}(kT.N)] H] \quad (\lambda \gg 2r). \end{aligned}\right\} \qquad (7.24)$$

In the presence of wind, i.e. the motion of the ion with respect to the earth's surface with the velocity u, the vector $\frac{uH}{c}$, is added to the vector E in the first of these equations, while the second equation remains unchanged.

The factor before E in the first equation represents the electrical conductivity of the ionized gas γ and for this reason the first equation in the presence of wind, may be written as follows:

$$j_{\perp} = \gamma \left(E + \frac{[uH_{\perp}]}{c} \right) \qquad (\lambda \ll 2r).$$

It is easy to see that if the term grad p' is rejected in equation 7.24, it then passes over into the Chapman formula for the drift current at $\varphi = 0$, but this is possible in the case when the partial pressure does not vary with the height. For the equilibrium state, however, the weight of ions in unit volume N_{mg} is balanced by the gradiant of partial pressure and, consequently $f_{\perp} = 0$.

In Tamm's opinion, the distribution of ions in the rear atmosphere differs so much from the equilibrium distribution that the difference N_{mg} - grad p' can hardly

be distinguished from the first term N_{mg}.

In this way, the theory of the drift currents in its modified form, exactly as with the dynamo theory, cannot be refuted and must be accepted for the explanation of the system of currents that cause the diurnal variations, but at the same time it still requires further developments, which Tamm himself admits.

It must be remarked that the Tamm theory enables us to explain the occurrence of the forenoon maximum of the horizontal component, which the Chapman theory could not do. More specifically, although the maximum number and weight of ions in N_{mg} should occur after noon, the maximum current may also occur before noon, since the maximum of the partial pressure of ions at the upper boundary of the upper ionosphere may occur after the maximum of the total ionization of this part of the ionosphere.

On comparing these two theories, it may be said that neither of them explains completely all observed facts and in some cases even contradict them. But still, owing to the successes that have recently been attained in the experimental study of the ionosphere, which have led to the establishment in them of periodic horizontal displacements and of the distance of a system of currents at height 100 - 110 km, and may be assumed that the dynamo affect undoubtedly plays a part in the formation of the system of currents causing the diurnal variations. This, however does not exclude the possibility of the simultaneous existence of the drift current effect as well, which may also participate in the formation of these currents.

The rapid development of our knowledge of the ionosphere will soon make it possible for us to solve the question which of these theories is correct and to eliminate the contradictions in them.

Section 3. Basic Propositions of the Theory of the Aurora and Magnetic Disturbance.

The close connection of the aurora with magnetic disturbances forces us to seek the causes that might be common to both phenomena. It has been remarked above that the magnetic activity that characterizes by which the magnetic disturbances are characterized is well correlated with the solar activity. For this reason all pres-

ent day theories of the magnetic disturbances are at the same time theories of the aurora as well, and start out with the assumption that their cause is energy radiated by the sun.

Since the energy of the magnetic field w in unit volume (cc) is expressed by the formula

$$w = \frac{H_T^2}{8\pi}.$$

then the increment of energy dw when the field strength H_T varies by the small quantity dH_T will obviously be

$$dw = \frac{H_T\, dH_T}{4\pi}.$$

Calculations by this formula give a value of $dw = 10^{-5}$ erg/cm^3 for the value of the mean density of the energy. The total value of this energy Δw, however is obtained by multiplying its density by the volume of the earth and the space around the earth, which gives $\Delta w = 3 \times 10^{22}$ ergs. The duration of a storm for which the energy density has been calculated was 28 hours, and therefore, by dividing the value of the total energy Δw by the number of seconds in 28 hours, we obtain the increment of the total energy of the magnetic field of the earth per seconds, equal to 3×10^{17} erg/sec. But the energy received by the earth from the sun amounts to 2×10^{24} erg/sec. In this way, only a negligible portion of the total energy radiated by the sun is required for the formation of the energy of the magnetic storms.

The reason why this part of the solar energy is converted into magnetic energy only at certain moments of time is connected with the phenomena in the active regions of the sun, such as spots, prominences, etc.

Observations of the sun show that these moments the intensity of the ultraviolet radiation is markedly increased, i.e. it is as though the temperature of the sun were increased. It is therefore assumed that during the periods of increased solar activity, the active regions, where the sunspots are, emit an intense ultraviolet

radiation and at the same time eject a large number of charged particles of both signs, electrons, ions, protons, and others.

Thus, for example, if 0.0001 part of the solar surface radiated as a black body at absolute temperature T=30.000°, then the solar constant would increase by 0.74 %, but the energy at wavelengths 3500, 4000, 5000 and 6000 A would increase respectively by 3.2, 1.7, 0.75 and 0.032 %. This additional energy would be entirely sufficient for the formation of the magnetic storms.

The repetition of magnetic disturbances at intervals of 27 days, equal to the period of rotation of the sun about its axis, furnished reason to assume that the radiation of the active regions takes place either in the form of a narrow beam rotating together with the solar disc, or in the form of a periodically acting wide beam. In this case it must be recognized that it is mainly the corpuscular radiation of the sun that participates in the formation of the magnetic storm and the aurora, although attempts have been made to attribute them mainly to the ultraviolet part of the solar spectrum. Such attempts, however have not led to positive results, and therefore the theory of the aurora based on the ultraviolet radiation of the active regions of the sun has at the present time lost its significance.

Section 4. The Motion of a Charge* in the Magnetic Field of the Dipole.

The fact that the aurora takes place mainly in the high latitudes close to the magnetic pole, and the direction of their streamers coincides with the direction of the lines of force from the earth's magnetic field, forces us to seek the source of the aurora in the motion of charges in the high layers of the atmosphere under the action of the earth's magnetic field.

Leaving, for the time being, the question open as to how and whence these charges appear, let us consider the motion of the charges in the earth's magnetic

*ZARYAD, charge, used in this section in the sense of "charged particle". Transla-

field, having the initial velocity **v**, under the assumption that the charges do not interact with each other. In this case, the problem reduces down to the consideration of the motion of one charge, an electron or a proton. The solution of this problem given by Störmer (Bibl.51) at the beginning of the present Century, consists in the solution of the differential equation of motion of a charge in the inhomogenous magnetic field of the earth, taken as the field of a dipole.

The equation of motion of a charged particle in the magnetic field has the form:

$$m\frac{d^2\mathbf{r}}{dt^2} = \frac{e}{c}\left[\frac{d\mathbf{r}}{dt}\mathbf{H}\right], \qquad (7.25)$$

where **e** is the charge of the particle, **m** its mass, and **r** the radius vector produced from the origin of coordinates.

If the origin of coordinates coincides with the center of the dipole, and the axis is directed along the axis of the dipole, the magnetic moment of which is **m**, then the field strength H is expressed as follows:

$$\mathbf{H} = \frac{3\mathbf{M}z}{r^5}\mathbf{r} - \frac{\mathbf{M}}{r^3}.$$

As a result of this, equation 7.25 takes the form

$$\frac{d^2\mathbf{r}}{dt^2} = \frac{Me}{mcr^3}\left[\frac{d\mathbf{r}}{dt}\left(3z\mathbf{r} - r^2\frac{\mathbf{z}}{z}\right)\right].$$

Introducing the new variable **ds = vdt**, where **v** is the velocity of motion of the charge and denoting **Me/mcv** = k^2, we obtain the following three equations scalar form:

$$\frac{1}{k^2}\cdot\frac{d^2x}{ds^2} = \frac{3z^2 - r^2}{r^5}\cdot\frac{dy}{ds} - \frac{3yz}{r^5}\cdot\frac{dz}{ds},$$

$$\frac{1}{k^2}\cdot\frac{d^2y}{ds^2} = \frac{3xz}{r^5}\cdot\frac{dz}{ds} - \frac{3z^2 - r^2}{r^5}\cdot\frac{dx}{ds},$$

$$\frac{1}{k^2}\cdot\frac{d^2z}{ds^2} = \frac{3z}{r^5}\left(y\frac{dx}{ds} - x\frac{dy}{ds}\right).$$

Multiply the first equation by **x** and the second by **y** and add them, term by term;

in exactly the same way, multiply this first equation by y and the second by x and subtract one from the other. As a result, the equations now take the following form:

$$\left.\begin{aligned}
\frac{1}{k^2}\left[x\frac{d^2x}{ds^2}+y\frac{d^2y}{ds^2}\right] &= \frac{3z^2-r^2}{r^5}\left[x\frac{dy}{ds}-y\frac{dx}{ds}\right],\\
\frac{1}{k^2}\left[y\frac{d^2x}{ds^2}-x\frac{d^2y}{ds^2}\right] &= \frac{3z^2-r^2}{r^5}\left[y\frac{dy}{ds}+x\frac{dx}{ds}\right]-\frac{3z}{r^5}\frac{dz}{ds}(x^2+y^2),\\
\frac{1}{k^2}\frac{d^2z}{ds^2} &= \frac{3z}{r^5}\left(y\frac{dx}{ds}-x\frac{dy}{ds}\right).
\end{aligned}\right\} \quad (7.26)$$

Project the vector r on the plane xy and denote its projection by R. Then

$$x = R\cos\varphi, \quad y = R\sin\varphi,$$

where φ is the angle between R and the x axis (Fig.108). By differentiating this expression with respect to s and performing the subtraction and addition of the derivatives, then, after multiplying them by y and x, and then by x and y, we obtain

$$x\frac{dy}{ds}-y\frac{dx}{ds}=R^2\frac{d\varphi}{ds}; \qquad x\frac{dx}{ds}+y\frac{dy}{ds}=R\frac{dR}{ds}.$$

On differentiating a second time with respect to s and performing the same operation of subtraction and addition, we shall have:

$$x\frac{d^2y}{ds^2}-y\frac{d^2x}{ds^2}=\frac{d}{ds}\left(R^2\frac{d\varphi}{ds}\right),$$

$$x\frac{d^2x}{ds^2}+y\frac{d^2y}{ds^2}=$$

$$=R\frac{d^2R}{ds^2}-R^2\left(\frac{d\varphi}{ds}\right)^2.$$

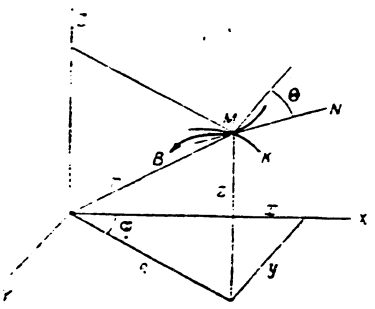

Fig.108.

The constant k may be taken as equal to unity since it is equivalent to all lengths being measured in units k.

On substituting these values of the sums and differences in equation 7.26 and

setting $k = 1$, we have:

$$\left.\begin{aligned} \frac{\partial^2 R}{\partial s^2} &= R\left(\frac{\partial \varphi}{\partial s}\right)^2 - R\frac{3z^2 - r^2}{r^5}\cdot\frac{\partial \tau}{ds}, \\ \frac{\partial}{\partial s}\left(R^2\frac{\partial \tau}{\partial s}\right) &= \frac{3zR^2}{r^5}\cdot\frac{\partial z}{\partial s} - R\frac{3z^2 - r^2}{r^5}\cdot\frac{\partial R}{\partial s}, \\ \frac{\partial^2 z}{\partial s^2} &= \frac{3z}{r^5}R^2\frac{\partial \tau}{\partial s}. \end{aligned}\right\} \qquad (7.27)$$

Remembering that $r^2 = R^2 + z^2$, it is easy to show that the following equations hold:

$$\frac{\partial}{\partial z}\left(\frac{R^2}{r^3}\right) = -\frac{3zR^2}{r^5}; \quad \frac{\partial}{\partial R}\left(\frac{R^2}{r^3}\right) = R\frac{3z^2 - r^2}{r^5}. \qquad (7.28)$$

The second equation of the preceding system is thus brought into the form:

$$\frac{\partial}{\partial s}\left(R^2\frac{\partial z}{\partial s}\right) = -\frac{\partial}{\partial z}\left(\frac{R^2}{r^3}\right)\frac{dz}{ds} - \frac{\partial}{\partial R}\left(\frac{R^2}{r^3}\right)\frac{\partial R}{\partial s}, \qquad (7.29)$$

or, after integration:

$$R^2\frac{\partial \tau}{\partial s} = -\left(2\gamma + \frac{R^2}{r^3}\right), \qquad (7.30)$$

where γ is the integration constant varying within the limits from $-\infty$ to $+\infty$.

On substituting the derivative $\partial\varphi/\partial s$ in the first and third equations by the system 7.27 we obtain from equation 7.29 and 7.30

$$\left.\begin{aligned} \frac{\partial^2 R}{\partial s^2} &= \left(\frac{2\gamma}{R} + \frac{R}{r^3}\right)\left(\frac{2\gamma}{R^2} + \frac{3R^2}{r^5} - \frac{1}{r^3}\right), \\ \frac{\partial^2 z}{\partial s^2} &= \left(\frac{2\gamma}{R} + \frac{R}{r^3}\right)\frac{3Rz}{r^5}. \end{aligned}\right\} \qquad (7.31)$$

Further, since ds is an element of the arc of the trajectory of the particle, then, by using its expression

$$ds^2 = dR^2 + dz^2 + R^2 d\tau^2,$$

we may write

$$1 - R^2\left(\frac{\partial \varphi}{\partial s}\right)^2 = \left(\frac{\partial R}{\partial s}\right)^2 + \left(\frac{\partial z}{\partial s}\right)^2,$$

or, substituting $\partial\varphi/\partial s$ from equation 7.30:

$$\left(\frac{\partial R}{\partial s}\right)^2 + \left(\frac{\partial z}{\partial s}\right)^2 = 1 - \left(\frac{2\gamma}{R} + \frac{R}{r^3}\right)^2 = Q. \qquad (7.32)$$

It is not hard to show that the righthand sides of equation 7.31 represent respectively the partial derivatives with respect to R and z of the function Q, separated into two, and thus the system of equations in equation 7.27 may be replaced by the system:

$$\frac{\partial^2 R}{\partial s^2} = \frac{1}{2}\cdot\frac{\partial Q}{\partial R}; \qquad \frac{\partial^2 z}{\partial s^2} = \frac{1}{2}\cdot\frac{\partial Q}{\partial z}; \qquad (7.33)$$

$$Q = \left(\frac{\partial R}{\partial s}\right)^2 + \left(\frac{\partial z}{\partial s}\right)^2. \qquad (7.34)$$

Consequently the integration of the system of equation 7.25 reduces down to the integration of equation 7.33 and 7.34.

Equation 7.33 shows that they determine the trajectory of the motion of a charge in the plane passing through the axis z and the radius vector R. The equation 7.34 however determines the rotation of this plane as a whole about the z axis.

Thus the problem of determining the motion of a charge in space reduces to two problems: determination of its motion in the plane zR and of the rotation of this plane about the z axis.

However, without solving the problem, but merely using the properties of the function Q, it is easy to find those regions of space in which the charge may move.

Since, according to equation 7.34, Q cannot assume negative values, the coordinates of the charge must always satisfy the inequality:

$$-1 \leqslant \frac{2\gamma}{R} + \frac{R}{r^3} \leqslant 1$$

all the equation

$$\frac{2\gamma}{R} + \frac{R}{r^2} = p, \qquad (7.35)$$

where p is a number less in absolute value than unity.

Consequently, for each value of the arbitrary constant γ_k there exists a region of space bounded on one side by the surface obtained by the roration of the curve defined by the equation

$$\frac{2\gamma_k}{R} - \frac{R}{r^2} = 1,$$

and by a similar surface defined by the equation:

$$\frac{2\gamma_k}{R} + \frac{R}{r^2} = -1.$$

this space, out of which the charge, at a given value of γ_k, cannot go, was denoted by Q_γ by Störmer.

The curve defined by equation 7.35 may be represented in polar coordinates if $R = r\cos\Phi$ is substituted in it and it is solved with respect to r, giving:

$$r = \frac{\gamma \pm \sqrt{\gamma^2 + p\cos^2\psi}}{p\cos\psi}. \qquad (7.36)$$

Störmer gives a complete analysis of these curves and the regions corresponding to them. He reached the following conclusions.

If $\gamma > 0$, not a single charge will be able to penetrate into the region near the dipole. If $\gamma < -1$, then for each γ there are two regions in which the charges may move, one of them laying outside the region of the dipole, so that the charges arriving from an infinite distance do not reach the dipole, while a second, a narrow region, does pass through the dipole. And, finally, if $-1 < \gamma < 0$, then both these regions merge into one and a charge coming from infinite distance may reach the dipole. Fig.109 shows a cross section of these regions in the meridional plane (the white parts of the skstch) and those regions to which the charges cannot penetrate

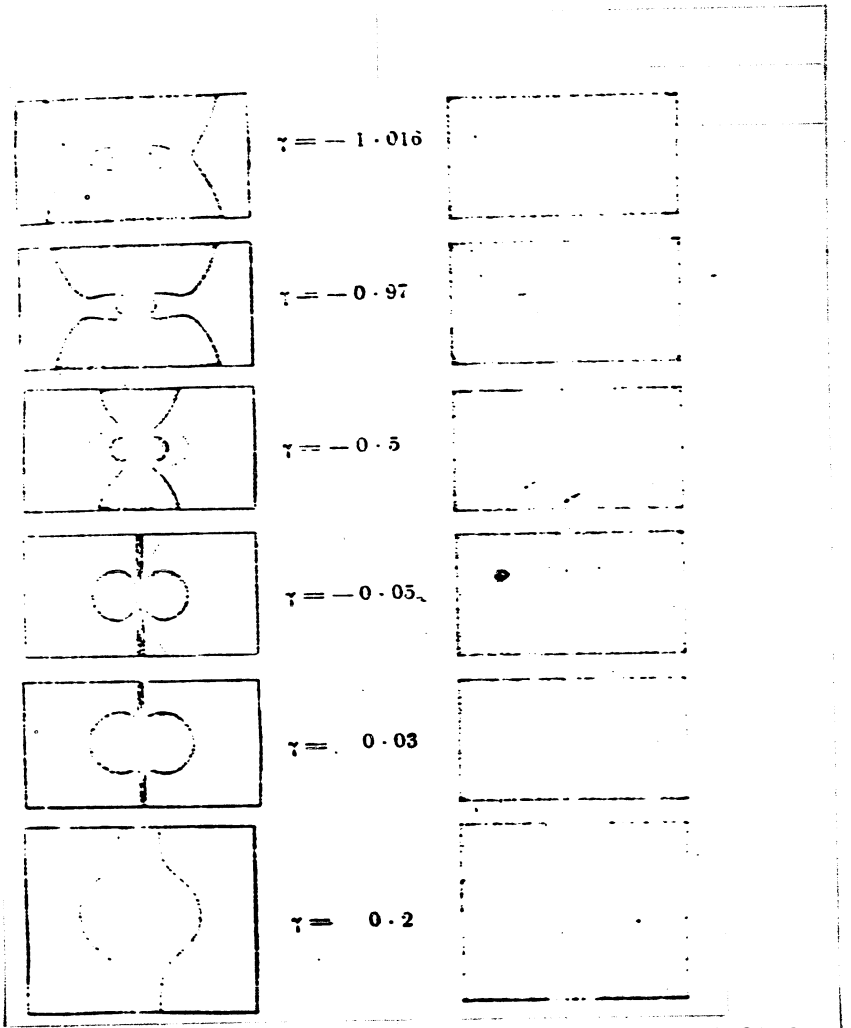

Fig.109. The Region Q in which Electrons may Move in the Magnetic Field of the Earth (not blackened).

(the black parts).

The following peculiarity attracts our attention: at $-1 < \gamma < 0$, there is a torus-like region around the dipole into which a charge from the outside cannot penetrate. The equation of the meridional curve of this torus will obviously be:

$$r = \frac{-\gamma_1 + \sqrt{\gamma_1^2 + \cos^2\varphi}}{\cos\varphi},$$

or, since r is always positive:

$$r = \frac{\cos^2\varphi}{\gamma_1 + \sqrt{\gamma_1^2 + \cos^2\varphi}}$$

where $\gamma_1 = -\gamma$.

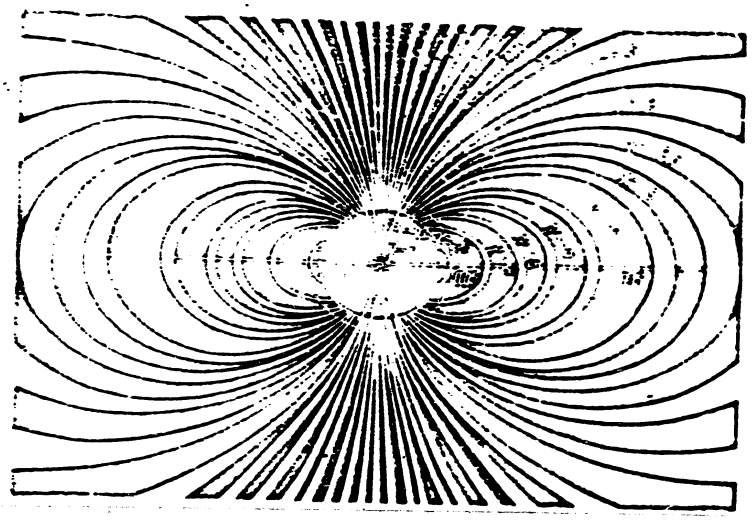

Fig.110. Regions Q in which Electrons can Move in the Earth's Magnetic Field Close to the Dipole (regions not blackened).

Fig.110 shows a meridional section of the regions near the dipole.

A consideration of these regions shows that for aurora only regions for which

$-1 < \gamma < 0$ are of importance, since by hypothesis, the charges come to the earth from the sun, i.e. from a practically infinite distance.

But for these values of γ the regions Q_γ make contact with the earth only in a limited part of its surface, close to the north and south magnetic poles, as may be seen from Fig.110. This part of the earth's surface is bounded by a circle which is the result of the section of the surface of the atmosphere by the surface of the torus defined by the equation

$$r = k\frac{\cos^2\frac{\psi}{2}}{1 \mp \sqrt{1 \mp \cos^2\frac{\psi}{2}}}. \qquad (7.37)$$

Here the factor $k = \sqrt{\frac{Me}{mcv}}$ is inserted so as to express r in centimeters.

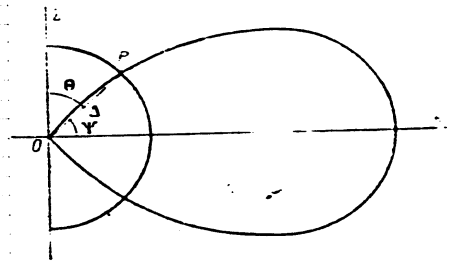

Fig.111.

Let this torus intersect the circumference of the upper layers of the atmosphere at the point P (Fig.111). Let us denote the distance from the center of the earth from the point P by Δ, and the angle zOP by θ. Then from equation 7.37 we obtain:

$$\Delta\left(1 \mp \sqrt{1 \mp \sin^2\theta}\right) = k\sin^2\theta,$$

whence

$$k^2\sin^2\theta - \Delta^2 - 2k\Delta = 0$$

and, consequently,

$$\sin\theta = \sqrt{\frac{2\Delta}{k} + \frac{\Delta^2}{k^2}}.$$

Since Δ is in practice small by comparison with k, we neglect the higher powers and obtain:

$$\sin\theta = \sqrt{\frac{2\Delta}{k}}.$$

Störmer found that the value of k for the beta-rays of radium is from 1.4×10^6

to 2.2 x 10^6 km; for the cathode rays, from 4.0 x 10^6 to 8.9 x 10^6 km; and for the alpha-rays, from 1.4 x 10^5 to 1.7 x 10^5 km.

Therefore the value of the angle Θ will be from 4 to 6° for beta-rays, from 2 to 4° for chathode rays, and from 16 to 19° for alpha rays, that is, the maximum deflection of particles from the magnetic pole amounts to 19°, while the maximum zone of the aurora is 23° from the pole and descends in strong magnetic storms to 30°, 40°, and even to 50°.

This contradiction between the theoretical conclusions and the observed facts may be circumvented if we assume the existence of a circular of a ring current in the plane of the magnetic equator at a distance of a few earth-radii. This current will reduce the strength of the earth's magnetic field and thereby will also reduce the value of k. In strong magnetic storms this current will increase, as a result of which the auroral zone expands.

Returning to the consideration of the system of equations 7.33 it must be noted that they are not integrated in simplest functions and therefore require approximate methods for their solution. However, without having recourse to their solution we may still draw certain conclusions as to the form of the trajectory in the meridional plane zR. Indeed, if the variable s is identified with time, then Q will represent a force (potential) function. For this reason, by plotting level lines Q at equal intervals on a plane, we can construct the direction of the forces acting on the charge. It is easy to see that when γ becomes negative, the force is equal to zero at the point $R = -\frac{1}{\gamma}$, z = 0, as well as on the levelled line Q = 1. The point $R = -\frac{1}{\gamma}$, z = 0 is a double point of the levelled line. Thanks to such a mechanical interpretation, the analysis of the trajectory is considerably simplified.

Between the trajectories in space and the trajectories in the meridional plane zR, there exists a simple geometrical relation. Let the point M (Fig.108) lie at the same time on the spatial trajectory B and on the trajectory K in the plane zR, and let the tangent N to the curve B at point M make the angle Θ with the meridional

plane. In that case, it follows from geometrical considerations that

$$\sin\theta = R\frac{d\varphi}{ds}, \quad \text{откуда} \quad \sin\theta = \frac{2\gamma}{R} + \frac{R}{\rho^2}.$$

Consequently, $\sin\Theta$ is the very same quantity that we denoted by p. Further, on introducing the function Q, we have:

$$\sin\theta = \mp\sqrt{1-Q}, \quad \cos\theta = \sqrt{Q}.$$

It follows from this that if the curve k intersects the levelled line $Q = 0$, then the tangent to the curve B will be perpendicular to the meridional plane. In this case, since there are always two levelled lines $Q = 0$, corresponding respectively to the values $p + 1$ and $p - 1$, and one of them lies within the line $Q = 1$, then, on intersecting the line $Q = 0$ inside $Q = 1$ the direction of the charge will coincide with the direction of the increasing values of Φ, while on its intersection outside the line $Q = 1$, the motion is opposite to the increase of the angle Φ.

However, to find the trajectories of the motion of the charge, besides these general deductions, Störmer also requires the solution of the system of equation 7.33 and 7.34. These solutions were performed by numerical integration in the form of infinite theories for r and φ.

Thus, for the values:- $-1 < \gamma < 0$ and $r < 0.2$, i.e. near the dipole, the following expressions were obtained for r and φ:

$$r = \frac{\cos^2\varphi}{2\gamma} + \frac{3}{8}\frac{1}{(2\gamma)^2}\left[\cos^{10}\varphi + \cos^{12}\varphi + \frac{15}{16}\cos^{14}\varphi + \frac{27}{32}\cos^{16}\varphi + \dots\right],$$

$$\varphi = -\frac{180}{\pi}\left[\frac{3}{16}r^2 + \frac{9\gamma}{32}r^3 + \frac{453\gamma^2}{1024}r^4 + \dots\right].$$

Fig. 112, which represents a wire model of the paths of charged particles near the dipole, shows the form of the curves corresponding to these equations for various values of γ.

Figs. 113 and 114 show a number of trajectories coming from infinity into the equatorial plane. Of these, Fig.113 relates to cases where $\gamma < -1$, while Fig.114

rolates to cases where γ ranges from -1 to -0.5. It will be seen that all trajectories for which $\gamma < -1$, lie outside the circle corresponding to $\gamma = -1$, while the trajectories having $0 > \gamma > -1$, penetrate inside this circle and consequently may approach the dipole.

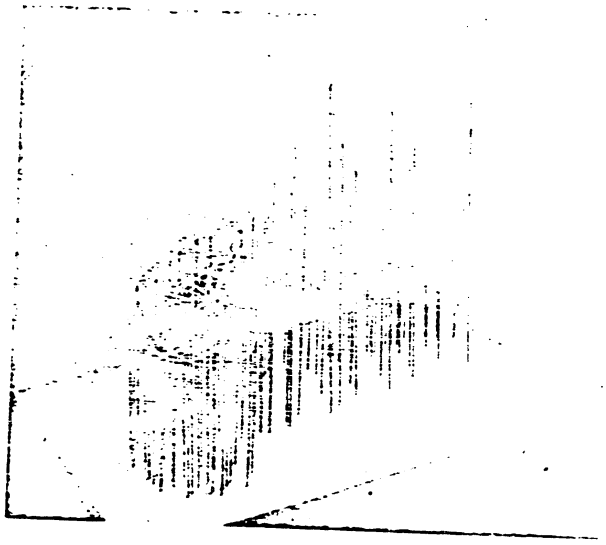

Fig.112. Wire Model of Trajectories of Electrons for Various Values of γ.

On applying the results of the mathematical conclusions directly to the aurora it may be said that, besides giving an explanation of the aurorals zone, they also explain the appearance of the aurora in the night hours, since the charges approaching the earth bend around it and leave the atmosphere on the opposite side from the sun. In exactly the same way the aurora observed during the daylight hours may be explained; they correspond to trajectories with small values of γ or with values from 0.93 to 1, which when the trajectories bend several times around the earth.

Further, the results of these conclusions form a correct interpretation of the radii form of the aurora. Indeed, in the radio forms of the aurora rapid and considerable variations in the intensity and position of the aurora are observed in a short time, and especially in the case of drapery. If we assume that the radial forms are formed by a narrow beam of charged particles emitted by the sun, then of the entire flux of particles on those which make a definite angle with the direction of the earth's magnetic axis ever reach the earth but since the magnetic axis, owing to the earth's rotation, varies its direction with respect to the trajectory of the particles, the direction of the rays also varies, i.e. some rays are extinguished, while others arise. But the arc-like of the aurora, still remain unexplained by this theory.

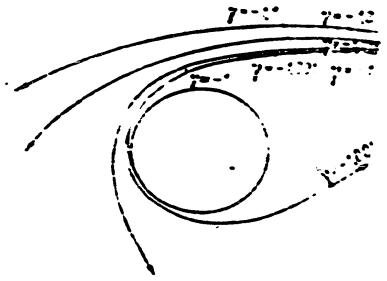

Fig.113. Trajectory of Electrons in the Equatorial Plane when $\gamma < 1$.

Such calculations, which threw much light on the formation of the aurora still could not cover the whole group of phenomena observed on their appearance. In exactly the same way they could not give even a functional connection between the motion of the charges near the dipole and the magnetic disturbances, since they indicated merely the qualitative aspect of this correlation, namely the moving charges must produce a magnetic field.

Since the mathematical conclusions are irreproachable, their failure to agree in many cases with the observed facts is explained by the unsoundness of the hypothesis that the electrons do not interact with each other during their motion.

However, taking into consideration that the trajectories of the charges calculated by Störmer do in many respects agree with the observed trajectories, and in addition, are confirmed by direct experiments made at various times by various in-

vestigators with cathode rays passing by a magnetized sphere, then, without rejecting the mathematical constructions and conclusions it is merely necessary to replace the fundamental hypothesis of a uniform, non-interacting beam of electrons by another, which would allow us to leave the mathematical calculations, on the whole,

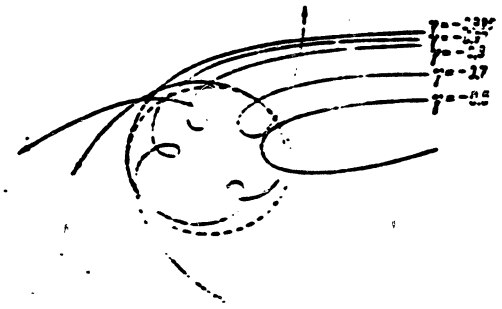

Fig.114. Trajectory of Electrons when $-1 < \gamma < -0.5$.

undisturbed, but considerably supplement it, instead of being contradicted, by the experimental data. The first of these attempts was made by Chapman and Ferraro, and the second by Alfven.

Section 5. The Theory of Chapman and Ferraro.

Störmer, at the same time as the solution of the problem of the motion of a charge in the magnetic field of the earth, attempted also to explain the formation of charges in the sphere of action of the earth's magnetic field. According to his hypothesis, these charges, of a single sign are ejected by the sun in the form of narrow beams, which, on reaching the earth's magnetic field, begins to move in accordance with the conclusions of this theory.

But elementary calculations show that a uniformly charge beam, owing to the electrostatic repulsion of the particles, must be scattered soon after its departure from the sun, and therefore could not under any conditions reach the earth.

To avoid this difficulty, Chapman and Ferraro (Bibl.58) constructed a theory on the hypothesis of the motion of a neutral beam.

According to their hypothesis, the sun emits a flux of particles in the form of a cylindrical beam consisting of equal numbers of positive charges of ions and of negative electrons.

Such a beam in diameter of the order of 50 earth-radii, is propagated from the sun with a velocity of 1000 km/sec, reaching the earth's orbit and intersecting it at an angle differing somewhat from 90°, since the particles, besides their longitudinal velocity, also have a transverse component equal to the linear velocity of rotation of the surface layers of the sun. The earth, moving on its orbit, may encounter on its path the lateral surface of the cylinder. Since the dimensions of the earth are small in comparison of the diameter of the cylinder, the lateral surface of the cylinder at the distance of a few earth radii may be taken as a plane, and in addition, the cylinder may be considered as motionless, and the earth itself as a dipole moving at a small angle in the direction of the plane, while the axis of the dipole may be taken as parallel to this plane. Since the medium bounded by the surface of the cylinder is conducting, it follows that when the dipole moves in it currents arise which will tend, by Lenz's law, to destroy the magnetic field of the dipole within the conductor.

In addition, between the dipole and the conducting medium, mechanical forces of repulsive character arise and tend to interrupt to motion of the dipole. Since the conducting medium is not a solid, but in its character rather recalls a gas, the plane surface, under the influence of these forces, will begin to be pressed in as it approaches the dipole, and under certain conditions, as shown by Chapman, a cavity surrounded by a concave conducting surface may be formed around the dipole.

Fig.115 shows the gradual distortion of the surface of the cylinder as it approaches the earth and the formation of a hollow space around the earth. In this case the vertex of the hollow, as will be seen from the figure, is turned towards

the sun, while on the night side there is free space.

If the ionized stream is an ideal conductor, then induction currents will arise only in a thin surface layer, and will completely screen the interior of the conductor from the penetration of the magnetic field of the dipole there. As a result of this, the magnetic lines of force in space between the dipole and the surface are more closely crowded together and thus cause an increase of the horizontal component in the equatorial plane, since the magnetic axis of the dipoles, by hypothesis, is parallel to the plane of the conductor.

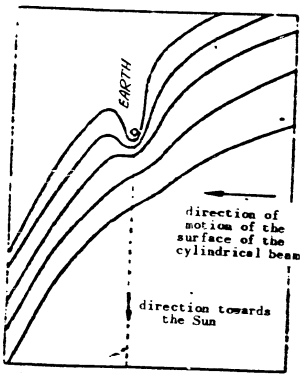

Fig.115. Section by the Equatorial Plane of the Cylindrical Beam of Charged Particles and Formation of a Hollow Base Around the Earth.

Such an increase in the horizontal component corresponds precisely to the first phase of a magnetic storm.

To explain the principal phase of a magnetic disturbance, during the time of which a fall in the horizontal component lasting several hours takes place, Chapman and Ferraro postulated that the free space on the night side might be closed off by ions detached from the surface of the cavity, and might form a closed ring current around the earth.

The possibility of the detachment of ions and their closure of the cavity is explained, according to Chapman and Ferraro by the following causes. On the surface of the cavity, owing to the deflecting action of the earth's magnetic field, the charges are separated, and, on one side, as is shown in Fig.116, the negative charges, or electrons, appear, and on the other side the positive charges or ions.

The Figure shows the direction of the magnetic axis of the earth (the plus sign) and the direction of the earth's movement, as well as the direction of motion of the

positive charges, i.e. the current. These charges produce an electric field on the night side of the earth, and under the influence of this field the ions and electrons will be detached from the surface. But a charge detached under the influence of the earth's magnetic fields would have to move along a circle whose radius is determined by the equation

$$r = \frac{mcv}{eH}, \qquad (7.38)$$

Consequently the flying charge might reach the opposite side of the cavity if the radius of its vortex motion r were of the same order as the width of the cavity, which, according to Chapman's calculation, amounts to several earth-radii. Calculations by equation 7.38 that such a value for the vortical motion at a velocity of the charge of the order of 1000 km/sec, could only be possessed by heavy charged particles, i.e. positive ions.

The closed mean current around the earth so formed must have the radius of a ring of the order of a few earth radii and would exist so long as the earth did not pass through the cylindrical beam. Since its diameter reaches 50 earth radii, while the orbital velocity of the earth's motion is about 20 km/sec , it follows that about 4-5 hours would be required to pass through the beam. During this period the horizontal component would reach a minimum, which corresponds to a maximum value of the current.

After passing through the beam, the ring current around the earth would be maintained, but would gradually decline owing to the scattering of the charged particles. In this way, qualitatively the Chapman theory would appear almost completely to explain one of the components of the magnetic variations, the aperiodic disturbed variation, as well as the 27-day cycle of these disturbances, since after a complete rotation of the sun, its surface, which radiates these beams, would again turn towards the earth and would again send out these beams to it.

The quantitative calculations made by Chapman and Ferraro, however, are con-

fined to the ideal schemes which give an order of magnitude corresponding to actuality.

The aurora and S_D disturbances are explained according to this theory by the possibility of detachment of electrons from the surface of the cavity, and, under proper conditions, of their motion along the lines of force from in the magnetic fields in the Arctic regions. From this point of view the electron beams must come not from the sun but from the surface of the cavity, as a result of which they are not able to become scattered during the short time interval, and the Störmers theory might therefore maintain its solidity.

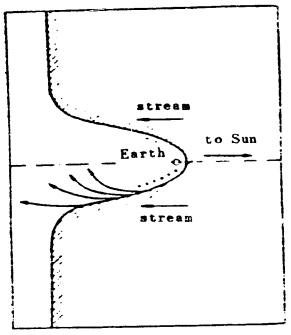

Fig.116. Formation of Positive and Negative Charged Particles on the Surface of the Hollow Space Around the Earth.

The irregular processes, i.e. the rapid variation of the magnetic field during the magnetic storm may be explained by the varying density of charged particles along the beam as well as by the different velocities of these charged particles.

The principal fault of the Chapman-Ferraro theory is the fact that it is not thoroughly worked out with respect to the aurora and its relation through the magnetic disturbances. While it gives an explanation of the sudden onset of a magnetic storm and the subsequent march of the aperiodic variations, it has not been able to explain either the local magnetic disturbances or the diurnal march of the magnetic variation.

A second fault of the theory is that it fails to give a quantitative relation between the magnetic field and the parameters of the neutral beam: its velocity, the density of the charges, the width of the beam, as well as the relation of these

beams with the aurora, and with other phenomena. For this reason it may be regarded as the first step in the construction in the theory of the magnetic disturbances and aurora, which is not contradicted by the basic observed facts and the fundamental laws of physics.

Section 6. Alfven's Theory.

The basic shortcomings of the Chapman-Ferraro theory, the difficulty of explaining the aurora and the local magnetic disturbances, as well as the diurnal march of magnetic disturbance, caused the search for somewhat different models of the motion of charged particles and the sun to the earth than those used by Chapman. Such a model was that of Alfven (Bibl.59) who in 1939 postulated that the magnetic field of the sun exerts a substantial influence on the motion of a neutral stream of particles, electrons and ions, while with Chapman and Ferraro, this motion takes place exclusively under the action of the initial impulse. This hypothesis allowed giving a more varied picture of the motion of particles of the magnetic field of the earth and the qualitative explanation of the origin of both their aurora and magnetic disturbances. On considering the question of the extent to which the sun's magnetic field influences the motion of the particles and how it affects their behavior near the earth, Alfven resolved the motion of particle in the magnetic field of the sun into the following four components:

1. Motion along a circle whose orbit is normal to the lines of the force of the magnetic field.

2. Oscillations along the lines of magnetic force about the plane of the solar equator, owing to the variations in magnetic field strength.

These oscillations along the lines of force result as a necessary consequence, from the theory of Störmer. Since no work is done when a charged particle moves in a magnetic field, the change in kinetic energy on oscillatory motion takes place on account of a change in the kinetic energy of the circular motion, as a result of which the particle is unable to deviate far from the equilibrium position, and there-

fore, according to Alfven, the oscillations are confined to the lower latitudes of the sun.

3. Deflection (in non-uniform drift u_1) of electrons towards the east and of positive ions towards the west, owing to the inhomogeneity of the magnetic field in a radial direction.

Such a deflection results in the production of space charges at the edges of the stream, negative on the eastern side and positive on the western, and, consequently, the appearance of an electric field perpendicular to the magnetic field.

4. The "drift" u_E of electrons in direction perpendicular to both and magnetic fields, owing to the simultaneous action of both fields.

The causes of this drift have already been considered in Section 3 of this Chapter. Owing to this drift, the particle moves in a radial direction from the sun to the earth.

All these motions are the result of the interaction of a charged particle with magnetic and electric fields and result from the general laws of electrodynamics. For this reason they represent nothing new, but their isolation from the general motion did allow the behavior of the electrons and ions near the earth to be followed up more graphically.

As they approach the earth, the particles use part of their energy, which originally, according the Alfven's hypothesis amounted to 10^8 ergs for electrons and somewhat less for the ions. These losses are due to the fact that the non-uniform drift u_1 transfers the particles into regions with a high potential electrical energy, where the presence of space charges retards the motion of the particles, and they lose parts of their velocity.

The drift u_E is part of the flux into the earth's magnetic field, the radius of action of which, by hypothesis, is of the same order as the width of the flux. Since the motion of the flux now will take place in an increasing magnetic field, the drift u_E due to the inhomogeneity of the magnetic field will become opposite to its

earlier drift and the particles, owing to this, will be transported into a region with a lower potential energy, and their magnetic energy will increase. The increase of kinetic energy will lead to an increase of the velocity perpendicular to the magnetic field which evidently becomes the same as the velocity parallel to the magnetic field, and the oscillations along the lines of force become small. As a result of this, the flux, seized by the earth's magnetic field begins to flatten out in its equatorial plane, and the non-uniform drift u_1 will tend to transport the positive ions around the earth in a westerly direction, and the electrons in an easterly direction. The ratio u_1/u_E is small outside the earth's magnetic field, but at a distance of about 5×10^9 cm from it, u_1 for the electrons is of the same order as u_E, while for the ions u_1 remains small. For this reason the electrons will be unable to pass into the "forbidden region" Q, established by Störmer, and must travel around it on the eastern side. The positive particles, however, freely penetrate it and form a space charge within it. As a result of electrostatic repulsion, the charged particles will be scattered, and the ions may move along the lines of force of the magnetic fields towards the magnetic poles. At the same time the electron beam E, bending around the earth and appearing on its night side, forms a negative space charge there which in turn will also be scattered and electrons are enabled to move along the lines of force to the magnetic poles, but this time from the night side. Fig.117 schematically shows the motion of ions and electrons in the beam near the earth.

Such is the mechanism of motion of charged particles from the sun to the earth according to Alfven hypothesis. But how are magnetic storms and the aurora explained according to this theory?

The aurora is the result of the penetrations of ions and electrons into the Arctic region as a result of their repulsion from the boundary of the forbidden region, and the auroral zone represents the projection onto the upper layers of the atmosphere of the lines of force emerging from the boundary of the forbidden region.

Since the equation of the lines of force of the dipole in polar coordinates has the form

$$r = R \sin^2 \vartheta,$$

where R is the value of the radius vector at $\Theta = 0^\circ$, then, if we take for R the radius of the boundary of the forbidden region, and for r and Θ the solar coordinates of the auroral zone which are known to us, we may find the radius R. If we consider $r \sim 6500$ km, and $\Theta \sim 20^\circ$, we shall have $R \sim 5 \times 10^9$ cm.

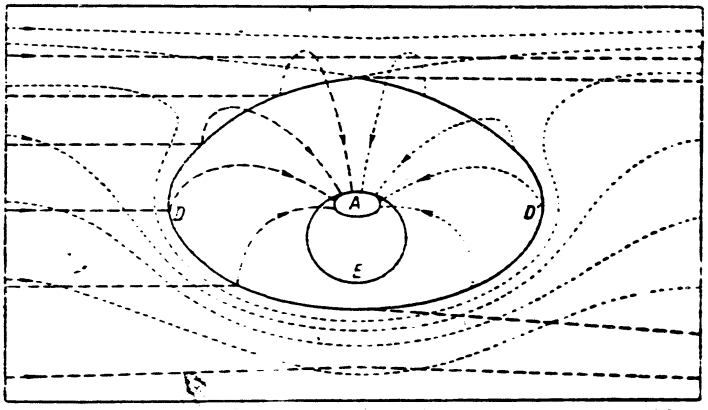

Fig.117. Scheme of Motion of the Positive Charged Particles (heavy dashed lines) and Negative Charged Particles (fine dashed lines) According to Alfven.

Boundary of Störmer's forbidden region Q; A auroral zone.

At such a distance the strength of the magnetic field of the earth in the plane of its equator, H = 70 Y. This value may correspond to that at which the electron begins to bend around the forbidden region Q.

Alfven then succeeded in explaining the various forms of the aurora including

the arcs which were not explained by the Störmer theory.

The principal phase of the magnetic storm is explained by the motion of electrons inside the forbidden region, which motion is equivalent to a closed current, while the local disturbances are explained by the appearance of currents in the auroral zone where, owing to the penetration of charged particles, an accumulation of positive charges takes place in the daylight part and of negative charges in the night part and in this way a potential difference is created.

The quantitative results of this theory depends on two unknown parameters, the energy and number of particles emitted by the sun. By a suitable selection of these quantities we may obtain a good agreement with the parameters known to us: the polar distance of the zone of the auroral zone, the energy of the electrons causing the aurora, the time of transit of the ionized beams of the sun to the earth, the duration of the initial phase of the magnetic storm, and the value of the current necessary to explain the storm.

However, in spite of the good agreement of certain observed parameters with their calculated values, the theory does not withstand serious criticism with respect to the formation of a flattened stream in the plane of the equator, which would be hardly possible. As a result of this, the formation of a narrow (linear) space charge along the forbidden region likewise becomes hardly possible. However, the fundamental factor of the Alfven theory is exactly the presence of such linear charges along the boundary of the forbidden regions. For this reason we must view the Alfven theory as one of the successive steps in the development of the ideas of the Störmer theory, and which, although it is not exempt from internal contradictions, still explains a wide circle of phenomena.